SIAM J. SCI. COMPUT.
Vol. 39, No. 4, pp. A1416–A1434

A BLOCK KRYLOV METHOD TO COMPUTE THE ACTION OF THE FRÉCHET DERIVATIVE OF A MATRIX FUNCTION ON A VECTOR WITH APPLICATIONS TO CONDITION NUMBER ESTIMATION*

PETER KANDOLF† AND SAMUEL D. RELTON‡

Abstract. We design a block Krylov method to compute the action of the Fréchet derivative of a matrix function on a vector using only matrix-vector products, i.e., the derivative of $f(A)b$ when A is subject to a perturbation in the direction E. The algorithm we derive is especially effective when the direction matrix E in the derivative is of low rank, while there are no such restrictions on A. Our results and experiments are focused mainly on Fréchet derivatives with rank 1 direction matrices. Our analysis applies to all functions with a power series expansion convergent on a subdomain of the complex plane which, in particular, includes the matrix exponential. We perform an a priori error analysis of our algorithm to obtain rigorous stopping criteria. Furthermore, we show how our algorithm can be used to estimate the 2-norm condition number of $f(A)b$ efficiently. Our numerical experiments show that our new algorithm for computing the action of a Fréchet derivative typically requires a small number of iterations to converge and (particularly for single and half precision accuracy) is significantly faster than alternative algorithms. When applied to condition number estimation, our experiments show that the resulting algorithm can detect ill-conditioned problems that are undetected by competing algorithms.

Key words. matrix function, matrix exponential, Fréchet derivative, condition number, Krylov subspace, block Krylov subspace

AMS subject classifications. 65F30, 65F60

DOI. 10.1137/16M1077969

1. Introduction.

Given a matrix function $f\colon \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$, one commonly needs to compute $f(A)b$ for a large, sparse matrix A and a vector $b \in \mathbb{C}^n$. For example, when $f(A) = \mathrm{e}^A$ is the matrix exponential, this computation arises when solving evolution equations with exponential integrators (e.g., [12], [13], [15]) or in network analysis (e.g., [6], [7]). In particular, [7] also uses the Fréchet derivative of the matrix exponential, which we introduce shortly.

To approximate $f(A)b$ one might use a Krylov subspace method. Since $f(A) = p(A)$ for some scalar polynomial $p(x)$, we construct the space

$$\mathcal{K}_m(A,b) = \operatorname{span}\left\{b, Ab, A^2b, \ldots, A^{m-1}b\right\} \tag{1.1}$$

and find a vector in this subspace which is close to $f(A)b$. We hope that, as the parameter m increases, the difference between the true value of $f(A)b$ and our approximation will decrease. More sophisticated methods that replace the polynomial

*Submitted to the journal's Methods and Algorithms for Scientific Computing section May 31, 2016; accepted for publication (in revised form) May 18, 2017; published electronically August 8, 2017.

http://www.siam.org/journals/sisc/39-4/M107796.html

Funding: The work of the first author was supported by a DOC Fellowship from the Austrian Academy of Science at the Department of Mathematics, University of Innsbruck, Austria. The work of the second author was supported by European Research Council Advanced Grant MATFUN (267526).

†Institut für Mathematik, Universität Innsbruck, Austria (peter.kandolf@uibk.ac.at, https://numerical-analysis.uibk.ac.at/p.kandolf).

‡Leeds Institute of Health Sciences, University of Leeds, Leeds, LS2 9NL, UK (s.d.relton@leeds.ac.uk, http://www.samrelton.com).

subspace (1.1) by a rational Krylov subspace, utilizing the inverse of shifted versions of A, have also been developed (see Güttel [8] for a recent survey). For the matrix exponentialthere are a number of alternatives to using a Krylov subspace, such as the Leja method [3] and methods based upon Taylor series approximation [1], which have been compared in a recent review paper [2].

In any numerical computation it is important to know the condition number of the problem: when the algorithm used is backward stable, we know that the relative error is approximately bounded above by the condition number times the unit roundoff. Therefore, being able to estimate the condition number efficiently is critical when determining how accurate a solution can be expected.

Before discussing the condition number of the $f(A)b$ problem, we must first define the Fréchet derivative of the matrix function f. The Fréchet derivative is an operator $L_f(A,\cdot)\colon \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$, which is linear in its second argument and for any matrix $E \in \mathbb{C}^{n\times n}$ satisfies

$$f(A+E) - f(A) = L_f(A,E) + o(\|E\|),$$

where $o(\|E\|)$ represents a remainder term that, when divided by $\|E\|$, tends to zero as $\|E\| \to 0$.

If we first consider the matrix function $f(A)$, without applying it to a vector, its relative condition number is closely related to the Fréchet derivative of the function. We have (see [10, Chap. 3])

$$\mathrm{cond}_{\mathrm{rel}}(f,A) := \lim_{\epsilon\to 0} \sup_{\|E\|\le\epsilon\|A\|} \frac{\|f(A+E)-f(A)\|}{\epsilon\|f(A)\|} = \max_{\|E\|=1} \frac{\|L_f(A,E)\|\|A\|}{\|f(A)\|}.$$

For the action of the matrix function, Deadman [5, p. 4] defines the relative condition number of $f(tA)b$. If one ignores the dependence on t by setting $t=1$, he defines

$$\mathrm{cond}(f,A,b) := \lim_{\epsilon\to 0} \sup_{\substack{\|E\|\le\epsilon\|A\| \\ \|\Delta b\|\le\epsilon\|b\|}} \frac{\|f(A+E)(b+\Delta b) - f(A)b\|}{\epsilon\|f(A)b\|} \tag{1.2}$$

for any induced matrix norm and goes on to show that

$$\mathrm{cond}(f,A,b) := \lim_{\epsilon\to 0} \sup_{\substack{\|E\|\le\epsilon\|A\| \\ \|\Delta b\|\le\epsilon\|b\|}} \frac{\|L_f(A,E)b + f(A)\Delta b + o(\|E\|) + o(\|\Delta b\|)\|}{\epsilon\|f(A)b\|}, \tag{1.3}$$

where the $o(\|E\|)$ and $o(\|\Delta b\|)$ terms disappear as $\epsilon \to 0$. In order to estimate this condition number efficiently we need to compute $L_f(A,E)b$, for multiple matrices E, in a highly efficient manner. The details of how to estimate this condition number are discussed further in section 5.

The primary goal of this paper is to give an efficient means of evaluating $L_f(A,E)b$ for E of rank 1 based upon block Krylov subspace methods. This can easily be extended to E of rank k, since $L_f(A,E)$ is linear in E. We also provide a priori error analysis for our algorithm in order to derive a rigorous stopping criterion. As a secondary goal we show how our algorithm can be applied to improve the computation of the condition number $\mathrm{cond}(f,A,b)$.

The paper is organized as follows. In section 2 we give some background information on the block Krylov method that we use throughout the rest of our analysis. Following this, in sections 3 and 4, we derive an algorithm to compute $L_f(A,E)b$ for a

rank 1 matrix E and perform our a priori error analysis. In section 5 we show how our algorithm can be used to estimate $\mathrm{cond}(f, A, b)$, before performing a battery of numerical experiments in section 6. Finally, in section 7 we give some concluding remarks.

2. Block Krylov method. The algorithm for computing $L_f(A, E)b$ that we introduce in section 3 is based upon the block Krylov decomposition; this section introduces the notation required throughout the rest of this work.

We define a block Krylov subspace, with block size p, as the set

$$\mathcal{K}(A, [x_1, \dots, x_p]) := \mathrm{span}\{x_1, \dots, x_p, Ax_1, \dots, Ax_p, A^2x_1, \dots, A^2x_p, \dots\}. \tag{2.1}$$

In practice we use a basis only for the span of the first $M = mp$ vectors, which we call $\mathcal{K}_m(A, [x_1, \dots, x_p])$. There are numerous variations of the block Arnoldi method available for computing a basis of this space, many of which can be found in [18, sect. 6.12]. In this work we have chosen to use the block Arnoldi–Ruhe algorithm introduced in [16] to compute a subspace with $p \geq 1$ starting vectors.

Algorithm 1 Block Arnoldi–Ruhe algorithm [16], [18].

Given a block size $p \geq 1$, the vectors $x_1, \dots, x_p$, and the size of the desired subspace $M = mp$, the following algorithm outputs an orthonormal basis, $\{v_j\}$ for $j = 1\!:\!M$, that spans the block Krylov subspace $\mathcal{K}_m(A, [x_1, \dots, x_p])$. When M is not a multiple of p, the algorithm returns an orthonormal basis which spans the first M vectors from the block Krylov subspace.

```
Compute p initial orthonormal vectors {v_i}_{i=1,...,p} that have the same span
  as {x_i}_{i=1,...,p} by extracting the columns of Q from QR = [x_1, ..., x_p].
for j = p : M + p − 1
    k = j − p + 1
    w = A v_k
    for i = 1: j
        h_{i,k} = w* v_i
        w = w − h_{i,k} v_i
    end for
    h_{j+1,k} = ‖w‖_2
    v_{j+1} = w / h_{j+1,k}
end for
```

Note that, for $p = 1$, the algorithm reverts to the standard Arnoldi process. One potential advantage of this algorithm over other variants is that the dimension of the subspace approximation M is not necessarily forced to be a multiple of p. This property can be of advantage in some applications, as it may require less computational effort to achieve convergence, but is not employed here. The following analysis assumes that M is a multiple of p.

When the subspace dimension M is a multiple of p, the algorithm can be rewritten to use level-3 BLAS routines, to take advantage of the induced block parallelism. To increase the stability of the algorithm we can also add a reorthogonalization step at the end of each iteration. In the rest of our analysis we will denote by m the number of steps and by M the dimension of the resulting block Krylov subspace.

For $M = mp$ we have analogues of the standard Arnoldi decomposition which are useful for our later analysis. Before giving these relationships, we need to define our notation. First, let us define U_i to be the $n \times p$ matrix with columns $v_{p(i-1)+1}, \dots, v_{pi}$

from Algorithm 1 and $V_m = [U_1, \ldots, U_m]$. Second, let $\bar{H}_m$ be the band Hessenberg matrix with nonzero entries h_{ij} from Algorithm 1. We split the matrix in the following fashion:

$$\bar{H}_m = \begin{bmatrix} H_m \\ 0_{p\times(m-1)p} \; H_{m+1,m} \end{bmatrix}, \qquad \text{with} \quad \begin{matrix} H_m & \in \mathbb{C}^{mp\times mp}, \\ H_{m+1,m} & \in \mathbb{C}^{p\times p}, \end{matrix}$$

and denote by J_m the final p columns of the identity matrix I_M. In this case we have the following analogues of the Arnoldi decomposition (see [18, p. 161]):

$$\begin{aligned} AV_m &= V_m H_m + U_{m+1} H_{m+1,m} J_m^T \\ &= V_{m+1} \bar{H}_m, \end{aligned} \tag{2.2}$$

$$V_m^T A V_m = H_m. \tag{2.3}$$

3. Algorithm derivation. In this section we derive the basic algorithm upon which the rest of our work builds. For the analysis in the next two sections we take the direction matrix $E = \eta y z^*$ in the Fréchet derivative to be of rank 1 with $\eta \in \mathbb{R}$ and $\|y\|_2 = \|z\|_2 = 1$. Furthermore, without loss of generality, we will assume that $\|b\|_2 = 1$. We can easily generalize the development to rank k matrices, as the Fréchet derivative is linear in its second argument. In fact, for a rank 2 matrix $E = \alpha_1 y_1 z_1^* + \alpha_2 y_2 z_2^*$ we have $L_f(A,E)b = \alpha_1 L_f(A, y_1 z_1^*)b + \alpha_2 L_f(A, y_2 z_2^*)b$, which we might expect to be in the block Krylov space $\mathcal{K}(A, [y_1, y_2, b])$. More generally, we would ideally like to use a block Krylov space with block size $k+1$ for an E of rank k. Analyzing this extension more formally is outside the scope of the current article but would make an interesting idea for future research.

We start by proving that the quantity that we wish to compute is in the Krylov subspace $\mathcal{K}_m(A, y)$ for some $m > 0$.

THEOREM 1. *Given A, $E \in \mathbb{C}^{n\times n}$ such that f is defined and Fréchet differentiable at A and such that $E = \eta y z^*$, then $L_f(A,E)b \in \mathcal{K}_m(A,y)$ for some $m > 0$, where $\mathcal{K}_m(A,y)$ is the Krylov subspace with starting vector y of dimension m.*

Proof. Assuming that the scalar function f is sufficiently differentiable, we can form a 2×2 block matrix to get

$$f\left(\begin{bmatrix} A & E \\ 0 & A \end{bmatrix}\right) = \begin{bmatrix} f(A) & L_f(A,E) \\ 0 & f(A) \end{bmatrix}.$$

We refer to [10, sect. 3.2] for the derivation of this formula. Therefore, we can multiply both sides by a vector from the right-hand side to obtain

$$\begin{aligned} f\left(\begin{bmatrix} A & E \\ 0 & A \end{bmatrix}\right) \begin{bmatrix} 0 \\ b \end{bmatrix} &= \begin{bmatrix} f(A) & L_f(A,E) \\ 0 & f(A) \end{bmatrix} \begin{bmatrix} 0 \\ b \end{bmatrix} \\ &= \begin{bmatrix} L_f(A,E)b \\ f(A)b \end{bmatrix}, \end{aligned} \tag{3.1}$$

so that the quantity we desire is in the upper half of the resulting vector.

Now for $X \in \mathbb{C}^{l\times l}$ we know that $f(X) = p(X)$ for some polynomial p of degree at most $l-1$ [10, Chap. 1], and therefore $L_f(A,E)b$ is in the span of the top half of the vectors formed by the matrix-vector products

$$\begin{bmatrix} A & E \\ 0 & A \end{bmatrix}^k \begin{bmatrix} 0 \\ b \end{bmatrix}$$

for $k = 0 : l-1$, where $l \leq 2n$. To complete the proof it remains to show that the top half of these vectors are in $\mathcal{K}_m(A, y)$, which we achieve via induction. The base case for $k = 0$ is obvious. For $k = 1$ and recalling that $E = \eta y z^*$, we have

$$\begin{bmatrix} A & \eta y z^* \\ 0 & A \end{bmatrix} \begin{bmatrix} 0 \\ b \end{bmatrix} = \begin{bmatrix} \eta y (z^* b) \\ Ab \end{bmatrix},$$

where the top half of the result is in $\mathcal{K}_1(A, y)$ and the bottom half is in $\mathcal{K}_1(A, b)$. For induction, suppose we have a vector

$$c_k = \begin{bmatrix} y_k \\ b_k \end{bmatrix},$$

where $y_k \in \mathcal{K}_k(A, y)$ and $b_k \in \mathcal{K}_k(A, b)$. Applying our 2×2 block matrix, we obtain

$$\begin{aligned} \begin{bmatrix} A & \eta y z^* \\ 0 & A \end{bmatrix} c_k &= \begin{bmatrix} A & \eta y z^* \\ 0 & A \end{bmatrix} \begin{bmatrix} y_k \\ b_k \end{bmatrix} \\ &= \begin{bmatrix} A y_k + \eta y (z^* b_k) \\ A b_k \end{bmatrix} \\ &=: \begin{bmatrix} y_{k+1} \\ b_{k+1} \end{bmatrix}, \end{aligned}$$

where $y_{k+1} \in \mathcal{K}_{k+1}(A, y)$ and $b_{k+1} \in \mathcal{K}_{k+1}(A, b)$, completing the proof. ∎

With this result we know that $L_f(A, yz^*)b \in \mathcal{K}_m(A, y)$, and our hope is that in practice a Krylov subspace with $m < n$ is required. Unfortunately, we have been unable to see how to approximate $L_f(A, yz^*)b$ solely from the space $\mathcal{K}_m(A, y)$. However, we will show that, by using the block Krylov subspace $\mathcal{K}_m(A, [y, b])$, we can compute an approximation to $L_f(A, yz^*)b$. Since $\mathcal{K}(A, b)$ is already required for computing $f(A)b$ with a Krylov method, this is not an unreasonable restriction.

When the scalar function f is analytic on and inside a contour Γ enclosing $\Lambda(A)$, the spectrum of A, we can represent $f(A)$ using the Cauchy integral formula (see [10, Def. 1.11])

$$f(A) := \frac{1}{2\pi i} \int_\Gamma f(\omega)(\omega I - A)^{-1} \mathrm{d}\omega. \tag{3.2}$$

Taking the Fréchet derivative of this representation using the chain rule [10, Prob. 3.8], we obtain

$$L_f(A, \eta y z^*) = \frac{\eta}{2\pi i} \int_\Gamma f(\omega)(\omega I - A)^{-1} y z^* (\omega I - A)^{-1} \mathrm{d}\omega. \tag{3.3}$$

For the next step, assume that we have a block Krylov subspace $\mathcal{K}_m(A, [y, b])$, so that we can form the approximation $(\omega I - A)^{-1} y \approx V_m (\omega I - H_m)^{-1} e_1$, and similarly $(\omega I - A)^{-1} b \approx V_m (\omega I - H_m)^{-1} \tilde{e}$, where e_1 is the first unit vector and $\tilde{e} = V_m^* b$. By employing these two approximations, we see that

$$\begin{aligned} L_f(A, \eta y z^*) b &\approx \frac{\eta}{2\pi i} \int_\Gamma f(\omega) V_m (\omega I - H_m)^{-1} e_1 z^* V_m (\omega I - H_m)^{-1} \tilde{e} \, \mathrm{d}\omega \\ &= \eta V_m L_f(H_m, e_1 (z^* V_m)) \tilde{e} \qquad (3.4) \\ &= \eta V_m L_f(H_m, e_1 \tilde{z}^*) \tilde{e}, \qquad (3.5) \\ &=: L^{(m)}, \end{aligned}$$

where $\tilde{z} = V_m^* z$ and $L^{(m)}$ is our approximation to $L_f(A, \eta y z^*)b$ from the Krylov subspace $\mathcal{K}_m(A,[y,b])$.

Overall, we have shown that if the block Krylov subspace $\mathcal{K}_m(A,[y,b])$ can be used to approximate both $(\omega I - A)^{-1}y$ and $(\omega I - A)^{-1}b$, then we can approximate $L^{(m)} \approx L_f(A, \eta y z^*)b$ by computing the Fréchet derivative at the (potentially much smaller) matrix H_m. We note that $H_m \in \mathbb{C}^{2m\times 2m}$, as we use a block Krylov subspace of block size 2.

To make this into an iterative algorithm we must now introduce a stopping criterion. We note that, in exact arithmetic, the algorithm will converge in at most $n/2$ iterations since we add two vectors in each iteration. If we want to compute the answer to some relative tolerance, tol, we can continue adding vectors into our Krylov subspace by increasing m until

$$\|L^{(m)} - L^{(m-1)}\|/\|L^{(m)}\| \leq \text{tol} \quad \text{or} \quad m = \text{floor}(n/2). \tag{3.6}$$

This results in the following basic algorithm.

Algorithm 2

Let A be an $n\times n$ matrix, and f a matrix function that is defined and Fréchet differentiable at A while being analytic on and inside a contour enclosing $\Lambda(A)$. Furthermore, suppose that y, z, and b are three vectors satisfying $\|y\|_2 = \|z\|_2 = \|b\|_2 = 1$. Finally, let $\eta \in \mathbb{R}$ and a tolerance tol > 0 be given. Then the following algorithm approximates $L_f(A, \eta y z^*)b$.

```
for m = 1: floor(n/2)
    Compute K_m(A,[y,b]) with the corresponding H_m and V_m.
    Compute L^(m) = η V_m L_f(H_m, e_1(z^* V_m)) ẽ, where ẽ = V_m^* b.
    if m > 1 and ||L^(m) − L^(m−1)||/||L^(m)|| ≤ tol,
        break out of loop
    end if
end for
return L^(m)
```

The above algorithm can be used to compute an approximation to $L_f(A,E)b$ but has one major issue: the rather crude stopping criterion can lead to early termination of the algorithm if any stagnation in the convergence is encountered. Therefore, our next section aims to design an a priori error bound so that (an upper bound on) the number of iterations required to achieve a given tolerance can be found analytically.

4. A priori error analysis. As mentioned above, the goal of this section is to improve the reliability of the stopping criterion in Algorithm 2 by performing an a priori error analysis. Our methodology is based upon that of Saad [17] for the matrix exponential, though we need to make a number of modifications to account for the Fréchet derivative. More specifically, we aim to show that our approximation to the Fréchet derivative at each step is equivalent to forming a bivariate polynomial in A and E. Once this has been achieved, we can use a bivariate power series as the remainder function and bound its norm. Since the size of this bound will decrease with m, we can find the number of iterations required to guarantee a given tolerance in our approximation.

During our analysis we will repeatedly use the fact that if a power series $f(x) = \sum_{i=0}^{\infty} a_i x^i$ has radius of convergence r, then for $A, E \in \mathbb{C}^{n\times n}$ with $\|A\| < r$ we have (see [10, Prob. 3.6])

$$L_f(A,E) = \sum_{i=1}^{\infty} a_i \sum_{j=1}^{i} A^{j-1} E A^{i-j}. \tag{4.1}$$

Our first step towards an a priori error estimate is to relate polynomials of the above form with corresponding polynomials in H_m, the block-Hessenberg matrix, which is the subject of the following lemma.

LEMMA 2. *Let A be any matrix and V_m, H_m the results of m steps of the block Arnoldi method applied to A with initial vectors $[y, b]$; also let $E = \eta y z^*$. Then for integers $0 \le i, j \le m-1$ we have*

$$A^i E A^j b = \eta V_m H_m^i e_1 z^* V_m H_m^j \tilde{e}, \quad \text{where } \tilde{e} = V_m^* b, \quad \text{and} \tag{4.2}$$

$$\sum_{s=0}^{j} \alpha_s \sum_{k=0}^{s} A^k E A^{s-k} b = \eta V_m \left(\sum_{s=0}^{j} \alpha_s \sum_{k=0}^{s} H_m^k e_1 z^* V_m H_m^{s-k} \right) \tilde{e}. \tag{4.3}$$

Proof. It suffices to prove (4.2), from which (4.3) follows easily. Writing out E in full, we have

$$A^i E A^j b = \eta (A^i y) z^* (A^j b).$$

Now, upon applying a result by Saad [17, Lem. 3.1], we have $A^i y = V_m H_m^i e_1$. Furthermore, since we know that $Ab = V_m H_m \tilde{e}$, where $\tilde{e} = V_m^* b$, we can apply a trivial modification of Saad's result to get $A^j b = V_m H_m^j \tilde{e}$. Making these substitutions, we obtain

$$A^i E A^j b = \eta V_m H_m^i e_1 z^* V_m H_m^j \tilde{e},$$

which completes the proof. ∎

Now, our plan is to approximate $L_f(A,E)b$ using a truncation of the power series expansion (4.1). As a corollary of our previous result, we show that truncating this series to order $m-1$ is equivalent to forming a polynomial using the basis of our Krylov subspace. In particular, let us define

$$s_{m-1}(A,E) = \sum_{i=1}^{m-1} a_i \sum_{j=1}^{i} A^{j-1} E A^{i-j} \tag{4.4}$$

to be a truncation of the power series (4.1).

COROLLARY 3. *Using the polynomial $s_{m-1}(x,e)$ defined in (4.4), we have $s_{m-1}(A, \eta y z^*) = \eta V_m s_{m-1}(H_m, e_1 z^* V_m)\tilde{e}$.*

Proof. This is a special case of (4.3) in Lemma 2. ∎

Therefore, the approximation $L_f(A, \eta y z^*)b \approx s_{m-1}(A,E)b$ is mathematically equivalent to utilizing the approximation $L_f(A, \eta y z^*)b \approx \eta V_m s_{m-1}(H_m, e_1 z^* V_m)\tilde{e}$. Using this result, we can now write an expression for the error which we will then bound. Before we do, let us define the remainder function as follows:

$$r_m(A,E) := L_f(A,E) - s_{m-1}(A,E).$$

This function quantifies how close our approximating polynomial $s_{m-1}(A,E)$ is to the actual Fréchet derivative $L_f(A,E)$. In the following lemma we will relate the remainder of $L_f(A,E)$ to the corresponding remainder of $\eta V_m L_f(H_m, e_1 z^* V_m) V_m^*$.

LEMMA 4. *Let A be any matrix and V_m, H_m the results of m steps of the block Arnoldi method applied to A with starting block $[y,b]$; also let $E = \eta y z^*$. Then*

$$L_f(A,E)b - \eta V_m L_f(H_m, e_1 z^* V_m)\tilde{e} = \eta\left[r_m(A, yz^*)b - V_m r_m(H_m, e_1 z^* V_m)\tilde{e}\right], \tag{4.5}$$

where the remainder $r_m(x,e)$ is defined above.

Proof. By rearranging the definition of the remainder polynomial, we have

$$\begin{aligned} L_f(A,E)b &= s_{m-1}(A,E)b + r_m(A,E)b \\ &= \eta V_m s_{m-1}(H_m, e_1 z^* V_m)\tilde{e} + r_m(A,E)b \qquad \text{by Corollary 3} \\ &= \eta V_m\left[L_f(H_m, e_1 z^* V_m)\tilde{e} - r_m(H_m, e_1 z^* V_m)\tilde{e}\right] + r_m(A,E)b. \end{aligned}$$

Now as $r_m(A,E)b = \eta\, r_m(A, yz^*)b$ we have

$$L_f(A,E) - \eta V_m L_f(H_m, e_1 z^* V_m)\tilde{e} = \eta\left[r_m(A, yz^*)b - V_m r_m(H_m, e_1 z^* V_m)\tilde{e}\right],$$

as required. ∎

Following this lemma, taking the norm of both sides, we have

$$\begin{aligned} \|L_f(A,E)b - \eta V_m L_f(H_m, e_1 z^* V_m)\tilde{e}\|_2 &= \|\eta[r_m(A, yz^*)b - V_m r_m(H_m, e_1 z^* V_m)\tilde{e}]\|_2 \\ &\le \eta\left(\|r_m(A, yz^*)b\|_2 + \|V_m r_m(H_m, e_1 z^* V_m)\tilde{e}\|_2\right) \\ &\le \eta\left(\|r_m(A, yz^*)\|_2 + \|r_m(H_m, e_1 z^* V_m)\|_2\right), \end{aligned} \tag{4.6}$$

where the final inequality is justified by the fact that $\|Xq\|_2 \le \|X\|_2$ for any X and q where $\|q\|_2 = 1$ and by the unitary invariance of the the 2-norm. Recall that (without loss of generality) at the start of section 3 we made the assumption that $\|b\|_2 = 1$ and therefore $\|\tilde{e}\|_2 = 1$.

With this result, we can now obtain a priori error bounds on our approximation which depend upon the size of the Krylov subspace. Using the polynomial $s_{m-1}(x,e)$ as defined in (4.4) at the mth iteration of the Krylov method, we obtain the remainder function

$$\begin{aligned} r_m(A,E) &= L_f(A,E) - s_{m-1}(A,E) \\ &= \sum_{i=m}^{\infty} a_i \sum_{j=1}^{i} A^{j-1} E A^{i-j}. \end{aligned} \tag{4.7}$$

Our next lemma bounds this remainder term from above using the power series $t_m(x) = \sum_{i=m}^{\infty} |a_i| i x^{i-1}$.

LEMMA 5. *Let f have a Taylor series convergent in a disc of radius r, let $\|A\|_2 < r$, and let f be Fréchet differentiable at A. Furthermore, let $E \in \mathbb{C}^{n\times n}$ be given. Then the remainder function* (4.7) *is bounded above by*

$$\|r_m(A,E)\|_2 \le \|E\|_2\, t_m(\|A\|_2).$$

Proof. We can see that

$$\begin{aligned}\|r_m(A,E)\|_2 &= \left\| \sum_{i=m}^{\infty} \sum_{j=1}^{i} a_i A^{j-1} E A^{i-j} \right\|_2 \\ &\leq \sum_{i=m}^{\infty} |a_i| \sum_{j=1}^{i} \|A\|_2^{j-1} \|E\|_2 \|A\|_2^{i-j} \\ &\leq \|E\|_2 \sum_{i=m}^{\infty} |a_i| i \|A\|_2^{i-1} \\ &= \|E\|_2 \, t_m(\|A\|_2). \qquad \square\end{aligned}$$

This leads to the following corollary.

COROLLARY 6. *For functions f and matrices A satisfying the criteria of Lemma 5, with $E = \eta y z^*$ and b given such that $\|b\|_2 = \|y\|_2 = \|z\|_2 = 1$, the error in approximating $L_f(A,E)b$ by $\eta V_m L_f(H_m, e_1 z^* V_m)\tilde{e}$ is bounded above by*

$$\|L_f(A,E)b - \eta V_m L_f(H_m, e_1 z^* V_m)\tilde{e}\|_2 \leq 2\eta t_m(\|A\|_2),$$

where t_m is defined just prior to Lemma 5.

Proof. Combine the upper bound (4.6) with Lemma 5, noting that $\|H_m\|_2 \leq \|A\|_2$ and that t_m is a monotonically increasing function. $\square$

To obtain more concrete bounds it is helpful to look at a specific function. For example, let us consider the matrix exponential. In this case we have the coefficients $a_i = 1/i!$ and hence $t_m(x) = \sum_{i=m}^{\infty} x^{i-1}/(i-1)!$, which is similar to the power series considered by Saad [17, Lem. 4.2 and Thm. 4.3]. Using Saad's result, we have the upper bound $t_m(x) \leq (x^{m-1}\mathrm{e}^x)/(m-1)!$, and therefore we have proven the following corollary.

COROLLARY 7. *When considering the matrix exponential in Corollary 6, we obtain the upper bound*

$$\|L_f(A,E)b - \eta V_m L_f(H_m, e_1 z^* V_m)\tilde{e}\|_2 \leq \frac{2\eta \|A\|_2^{m-1} \mathrm{e}^{\|A\|_2}}{(m-1)!}.$$

Proof. See above. $\square$

The analysis in this section has allowed us to use a relative a priori error bound to terminate our iterative algorithm, as opposed to simply taking the relative difference between iterates in Algorithm 2. The resulting algorithm is given below.

5. Application to condition number estimation. In this section we use our new algorithm for computing $L_f(A,E)b$ to design an algorithm for estimating the condition number of $f(A)b$, building upon work by Deadman [5]. In particular, Deadman proves the following bounds on $\mathrm{cond}(f,A,b)$, which we defined in (1.3).

Algorithm 3

Let A be an $n \times n$ matrix, and f a matrix function that is defined and Fréchet differentiable at A while being analytic on and inside a contour enclosing $\Lambda(A)$. Furthermore, suppose that y, z, and b are three vectors satisfying $\|y\|_2 = \|z\|_2 = \|b\|_2 = 1$. Finally, let $\eta \in \mathbb{R}$ and tol > 0 be given. Then the following algorithm attempts to approximate $L_f(A, \eta yz^*)b$ within a relative tolerance tol.

```
for m = 1: floor(n/2)
    Compute K_m(A, [y, b]) with the corresponding H_m and V_m.
    Compute L^(m) = η V_m L_f(H_m, e_1(z* V_m)) ẽ, where ẽ = [I_M 0] V_m* b.
    Set α equal to the a priori error estimate given by Corollary 6 or 7.
    α = α / ||L^(m)||_2        % Make the bound relative
    if m > 1 and α < tol,
        break out of loop
    end if
end for
return L^(m)
```

LEMMA 8. *Given $A \in \mathbb{C}^{n\times n}$, $b \in \mathbb{C}^n$, an induced matrix norm $\|\cdot\|$, and a matrix function f which is defined and Fréchet differentiable at A, the condition number* $\mathrm{cond}(f, A, b)$ *satisfies the following bounds:*

$$\frac{1}{\|f(A)b\|} \max\left(\|A\| \max_{\|E\|=1} \|L_f(A,E)b\|, \|f(A)\|\|b\|\right) \le \mathrm{cond}(f,A,b) \le \frac{1}{\|f(A)b\|}\left(\|A\| \max_{\|E\|=1} \|L_f(A,E)b\| + \|f(A)\|\|b\|\right).$$

Proof. For a proof, see [5, p. 5]. ∎

Deadman then proceeds to work in the 1-norm and shows that

$$\mathrm{cond}(f,A,b) \approx \left(2\sqrt{n}\gamma\|A\|_1 + \|f(A)\|_1\|b\|_1\right)/\|f(A)b\|_1, \tag{5.1}$$

where $\|A\|_1$ and $\|f(A)\|_1$ can be estimated using the 1-norm estimation algorithm by Higham and Tisseur [11]. The quantity γ is a 2-norm estimate of the matrix $K_f(A,b) \in \mathbb{C}^{n\times n^2}$ (not to be confused with the Krylov subspace), which is a linear operator such that $K_f(A,b)\,\mathrm{vec}(E) = \mathrm{vec}(L_f(A,E)b)$ for any matrix $E \in \mathbb{C}^{n\times n}$ [5, p. 7]. It is simple to show that $K_f(A,b) = (b \otimes I_n)K_f(A)$, where $\otimes$ represents the Kronecker product and $K_f(A)$ is the Kronecker form of the Fréchet derivative [10, Chap. 3].

Instead of following Deadman by approximating the 1-norm condition number, we choose to estimate the 2-norm condition number, for which this approximation becomes

$$\mathrm{cond}(f,A,b) \approx (2\gamma\|A\|_2 + \|f(A)\|_2\|b\|_2)/\|f(A)b\|_2. \tag{5.2}$$

Note that this avoids introducing the $\sqrt{n}$ factor in the first term which, for large sparse matrices, could potentially dominate the approximation. We can then estimate $\|f(A)\|_2$ and $\|A\|_2$ using a power method. This leaves only the estimation of γ.

To estimate γ we first need to introduce the adjoint of a Fréchet derivative, $L_f^\star(A,E)$. For simplicity we will focus on matrix functions with real power series (such as the matrix exponential) for which we know that $L_f^\star(A,E) = L_f(A^*,E)$. For further details on the adjoint, see Higham [10, p. 66], for example.

Algorithm 4 Condition number estimate; cf. [5, Alg. 3.3].

For the function f, $A \in \mathbb{C}^{n\times n}$, and $b \in \mathbb{C}^n$ this algorithm computes an estimate $\gamma \le \|K_f(A,b)\|_2$.

1 Choose a nonzero starting vector $y_0 \in \mathbb{C}^n$.
2 it_max $= 10$ % Choose the maximum number of iterations.
3 for $k = 0 : \text{it\_max}$
4 $\quad y_{k+1} = L_f(A, L_f^\star(A, y_k b^*))b$
5 $\quad \gamma_{k+1} = \sqrt{\|y_{k+1}\|_2}$
6 $\quad$ if $|\gamma_{k+1} - \gamma_k| < 0.1\gamma_{k+1}$, $\gamma = \gamma_{k+1}$, quit, end
7 $\quad y_{k+1} = y_{k+1}/\|y_{k+1}\|_2$
8 end
9 $y = y_{k+1}$

Using the adjoint, the value $\gamma \approx \|K_f(A,b)\|_2$ is estimated from the following algorithm, given by Deadman [5].

The major cost of this algorithm is computing the derivative on line 4, which can be computed by extracting the top n elements of the vector

$$f\left(\begin{bmatrix} A & L_f^\star(A, y_k b^*) \\ 0 & A \end{bmatrix}\right)\begin{bmatrix} 0 \\ b \end{bmatrix}. \tag{5.3}$$

Let us denote the 2×2 block matrix in the above equation by X, and the vector by v. Furthermore, we denote by $\texttt{fAmv}(X, v)$ a method that computes $f(X)v$ using only matrix-vector products. This quantity could be computed using a Krylov method or any other method requiring only matrix-vector products. However, such a method will have to deal with the inner subproblem of computing matrix-vector products with $L_f^\star(A, y_k b^*)$, found in the top-right block of the matrix in (5.3). Following the approach taken by Deadman, since $L_f^\star(A, y_k b^*) = L_f(A^*, y_k b^*)$ for functions with real power series expansions, we could compute $L_f^\star(A, y_k b^*)v$ as the top n elements of

$$f\left(\begin{bmatrix} A^* & y_k b^* \\ 0 & A^* \end{bmatrix}\right)\begin{bmatrix} 0 \\ v \end{bmatrix}.$$

Now, since the direction term in the Fréchet derivative is of rank 1, our new algorithm can compute this inner subproblem in an extremely efficient manner. Note that, since the condition number does not generally need to be known to high precision, our new algorithm can be run with a low tolerance for even greater efficiency. Before we give the overall algorithm, we first define the following code fragment for efficiently computing the quantity given by (5.3).

CODE FRAGMENT 9 ($w = Xu$). *This code computes the matrix-vector product*

$$\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} A & L_f^\star(A, y_k b^*) \\ 0 & A \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix},$$

where the direction term in the Fréchet derivative is of rank 1 *and computed by Algorithm* 2 *or* 3.

1 $w_1 = Au_1 + L_f(A^*, y_k b^*)u_2$
2 $w_2 = Au_2$
3 return $w = [w_1^T, w_2^T]^T$

Algorithm 5

Given $A \in \mathbb{C}^{n\times n}$, $y_k \in \mathbb{C}^n$ and $b \in \mathbb{C}^n$ with $\|b\|_2 = \|y_k\|_2 = 1$. The user-supplied routine `fAmv` for which $\texttt{fAmv}(X,v) = f(X)v$ uses only matrix–vector products $w = Xu$ is evaluated by Code Fragment 9. The algorithm computes $x = L_f(A, L_f^\star(A, yb^*))b$.

1 $v = [0, b^T]^T$
2 $x = \texttt{fAmv}(X,v)$ computes $f(X)v$ by employing Code Fragment 9.
3 $x = x(1\colon n)$
4 return x

Note: The matrix $X \in \mathbb{C}^{2n\times 2n}$ denotes the 2×2 block matrix in (5.3).

Altogether, this leads to Algorithm 5 which is used to compute $L_f(A, L_f^\star(A, y_k b^*))b$ on line 4 of Algorithm 4.

By combining the bound on the condition number (5.2) with the estimate of γ in Algorithm 4, where we replace line 4 with Algorithm 5, we obtain the following algorithm for estimating the 2-norm condition number.

Algorithm 6

Given f, A, and b satisfying the criteria of Algorithms 2 and 3, and assuming that an algorithm for computing $f(A)b$ is available, the following algorithm computes $f(A)b$ and estimates its condition number $\kappa \approx \text{cond}(f, A, b)$ in the 2-norm.

1 Compute $f(A)b$ using any algorithm.
2 Estimate $\|A\|_2$ and $\|f(A)\|_2$ using the power method.
3 Compute γ using Algorithm 4, replacing line 4 with Algorithm 5.
4 $\kappa = (2\gamma\|A\|_1 + \|f(A)\|_2\|b\|_2)/\|f(A)b\|_2$.

6. Numerical experiments. In this section we will perform a number of numerical experiments to test the accuracy and efficiency of our new algorithms. All experiments were performed using MATLAB 2015a on a Linux machine (specifically the 64-bit variant—glnxa64), and all timings use only one processor which we achieved by using the option `-singleCompThread`. We also deactivated the Java virtual machine (and hence ran MATLAB without a GUI) by using `-nojvm`, to further increase the reliability of the results. The Python code is run using the Anaconda distribution of Python 2.7, with the libraries scipy 0.17.0 and numpy 1.10.4, linked to Intel MKL BLAS.

We use the matrix exponential (almost) exclusively throughout our tests, since this allows us to test the performance of our algorithm using `expmv` [1], which has implementations in both MATLAB and Python. We can use `expmv` to compute $L_{\exp}(A,E)b$ using the 2×2 block formula from (3.1). We will refer to `expmv` used within the block formula for the Fréchet derivative as the "block algorithm."

Throughout our experiments the direction matrix E will be rank-1 such that $E = yz^*$, where y and z have elements sampled from a random normal $N(0,1)$ distribution.

We will also need to compute $L_{\exp}(A,E)b$ in a highly accurate manner, so that we can compute the relative error achieved by our algorithm. We compute this by forming the 2×2 block formula (3.1) in variable-precision arithmetic from the MATLAB Symbolic Toolbox. We compute the function $f(X)$ to high precision by taking an eigendecomposition $VDV^{-1} = X + \Delta X$ using 200 digits, where $\|\Delta X\|_2 = 10^{-100}$ is used to ensure that, with probability 1, the matrix is diagonalizable. We can then

form $Vf(D)V^{-1}$ and round back to double precision floating point.

Our first two experiments are designed to investigate the behavior of Algorithms 2 and 3. We compare the relative errors achieved by the block algorithm and Algorithm 2 when aiming for half, single, or double precision accuracy over a range of test matrices. We also compare the amount of work performed by each algorithm.

Our second experiment shows the convergence of our new algorithm as m, the number of block Krylov iterations, increases. This allows us to compare the effectiveness of the two stopping criteria in Algorithms 2 and 3. We also include a convergence profile using the matrix function $\varphi(A) = A^{-1}(\mathrm{e}^A - I)$ in this experiment.

The third experiment illustrates the performance of our new algorithm on large sparse matrices. We also illustrate the storage advantages of our algorithm in comparison to the block algorithm.

Finally, in our fourth experiment, we investigate the application of our algorithm to computing a bound on the condition number of $\mathrm{e}^A b$ for several different matrices.

Experiment 1 (behavior in floating point arithmetic). This experiment is designed to test the behavior of the method in floating point arithmetic. To do this, we will check that the algorithms attain the prescribed relative error. In particular we use the tolerances *half, single*, and *double,* which in IEEE floating point arithmetic correspond to 2^{-11}, 2^{-24}, and 2^{-53}, respectively.

The test matrices used in this experiment are taken from the Matrix Computation Toolbox [9], where we set their dimension to $n = 100$ and compute $L_{\exp}(A, yz^*)b$ for y, z, and b with elements sampled from a normal $N(0, 1)$ distribution. For reference we also compute $L_{\exp}(A, yz^*)b$ by applying `expmv` to the 2×2 block matrix shown in (3.1), i.e., the block algorithm. We compute the exact solution, from which we calculate the relative error obtained, using variable precision arithmetic as described previously.

In an attempt to control the, sometimes extremely large, condition numbers of the matrix functions, we rescaled all the matrices to have unit 2-norm. Nevertheless, several of the matrices were still too ill-conditioned to yield reasonable results with either method. The excluded matrices were those with indices $[4, 7, 12, 17, 18, 29, 35, 40]$ from the Matrix Computation Toolbox.

In Figure 6.1 we see the relative error obtained by each algorithm. The results show that, for each test matrix, both algorithms obtain a relative error close to the desired accuracy. In particular this shows that our new algorithm appears to behave in a forwards stable manner.

In Figure 6.2 we show the amount of work required by each algorithm. We define one unit of work to be the effort (in flops) required to compute one matrix-vector product with an $n \times n$ matrix. Therefore, if our new algorithm is using an $m \times m$ matrix on the current iteration, then each matrix-vector product contributes m^2/n^2 units of work. We see that the block algorithm performs a similar amount of work regardless of the accuracy desired: they are barely distinguishable on the graph since all three lines overlap one another. Meanwhile, our new algorithm is often cheaper when aiming for double precision accuracy, and over a factor of 10 times cheaper for single and half precision accuracy.

In Figure 6.3 we see how the savings in work described above translates to savings in time. In order to obtain reliable timings we ran each code 500 times and display the mean time. We see that Algorithm 2 is faster when the desired accuracy is half or single precision. However, with a few exceptions, it is slower when we desire double precision accuracy. This is partially due to the fact that inside our implementation we

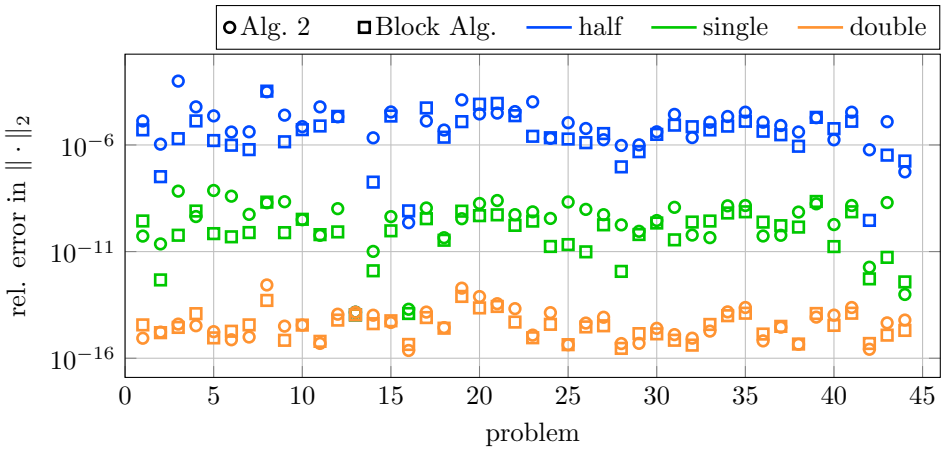

FIG. 6.1. *Relative errors obtained for the computation of $L_{\exp}(A,E)b$ by the new algorithm and the 2×2 block form when aiming for half, single, and double precision accuracy.*

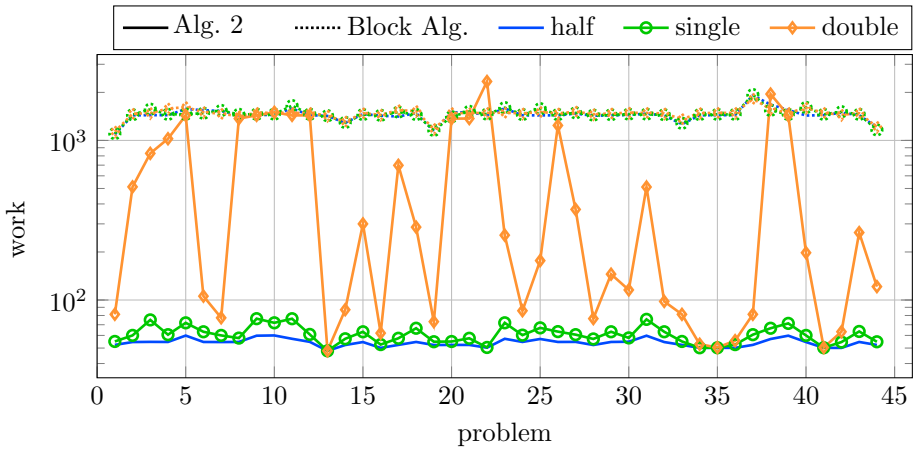

FIG. 6.2. *Units of work required to compute $L_{\exp}(A,E)b$ for the examples in Figure 6.1. Each unit of work is equivalent to one matrix-vector product with an $n \times n$ matrix.*

evaluate the stopping criterion, (3.6), at the end of every iteration. We may be able to improve the runtime by, for example, checking the stopping criterion less frequently.

Experiment 2 (convergence profile). In our second experiment we illustrate the convergence behavior of Algorithms 2 and 3 with respect to the degree m of the block Arnoldi process; see Figure 6.4. This will allow us to compare the two different stopping criteria used in Algorithms 2 and 3.

For the convergence profile we have chosen the matrix A to be the `poisson2D` matrix from the University of Florida Sparse Matrix Collection [4]. The vectors b, y, and z are chosen randomly with elements sampled from a normal $N(0,1)$ distribution. As before, we begin by computing $L_{\exp}(A, yz^*)b$.

For the computation shown in Figure 6.4 the stopping criteria are ignored, so that we continue iterating until $m = \text{floor}(n/2)$. The exact solution is computed in variable

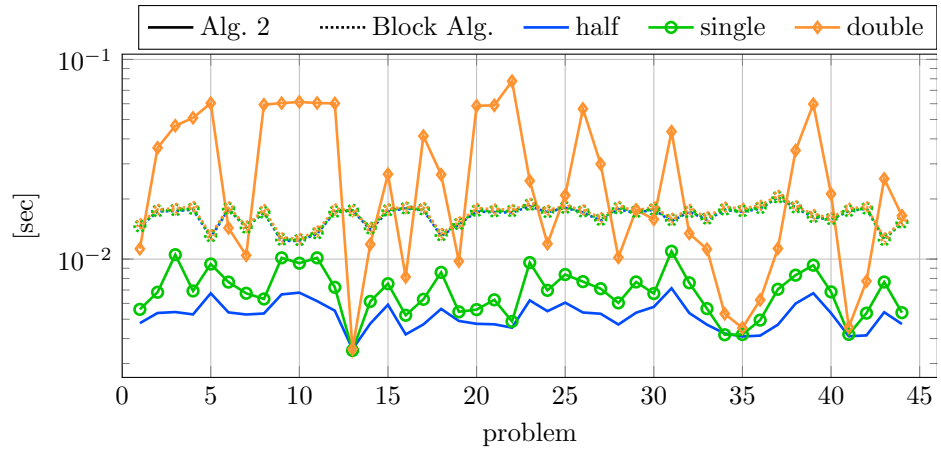

FIG. 6.3. *Wall time (in seconds) required for the computation of the examples in Figure* 6.1.

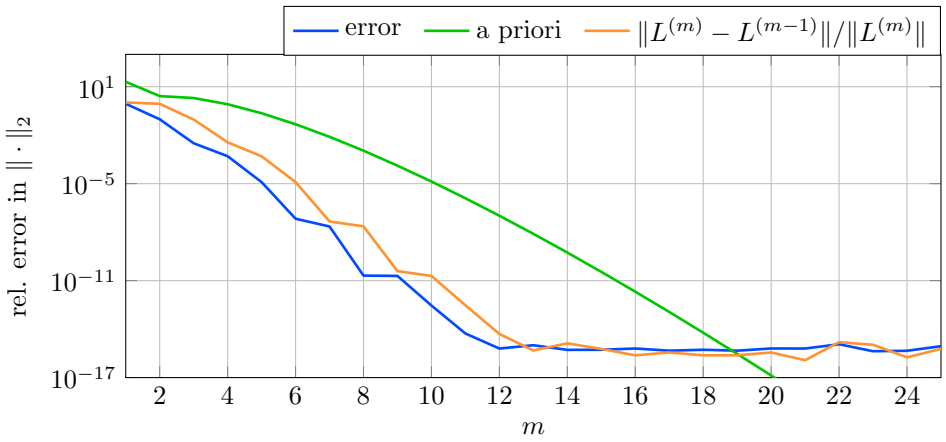

FIG. 6.4. *Convergence behavior of Algorithm* 2 *and the a priori error estimate as well as the heuristic early termination criterion for the matrix function* $\exp(A)$. *The error is measured in the* 2*-norm.*

precision arithmetic with 200 digits in the manner described at the beginning of this section. The graphs are cropped at $m = 25$ as, by this point, full double precision accuracy has already been achieved.

As we might expect, we see that the a priori estimate overestimates the number of iterations required to compute the Fréchet derivative at the desired accuracy but certainly provides an upper bound on the number of iterations required. On the other hand, we can see that the heuristic stopping criterion, $\|L^{(m)} - L^{(m-1)}\|/\|L^{(m)}\|$, captures the rate of convergence well.

In our second convergence profile we switch to computing $L_\varphi(A, yz^*)b$, where $\varphi(A) = A^{-1}(\exp(A) - I)$, using the same matrix and vectors as above. The matrix function $\varphi(x)$ is used extensively in exponential integrators [14]. In order to use our a priori error bound we must tailor Lemma 5, and the corollaries following it, for use with the $\varphi(x)$ function. In particular, this function has the Taylor series $\varphi(x) = \sum_{i=0}^{\infty} x^i/(i+1)!$. Therefore the a priori error (defined just before Lemma 5)

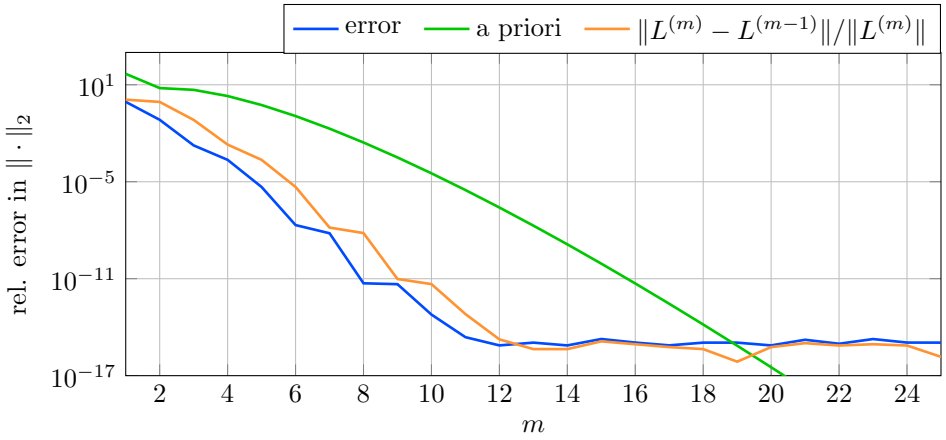

FIG. 6.5. *Convergence behavior of Algorithm* 2 *and the a priori error estimate as well as the heuristic early termination criterion for the matrix function* $\varphi(A)$. *The error is measured in the* 2*-norm.*

can be bounded as follows:

$$\begin{aligned} t_m(x) &= \sum_{i=m}^{\infty} \frac{i}{(i+1)!} x^{i-1} \\ &= \sum_{i=m}^{\infty} \frac{x^{i-1}}{(i-1)!} \left(\frac{1}{i+1} \right) \\ &\leq \frac{1}{m+1} \sum_{i=m}^{\infty} \frac{x^{i-1}}{(i-1)!} \\ &\leq \frac{1}{m+1} \left(\frac{2\eta \|A\|_2^{m-1} \mathrm{e}^{\|A\|_2}}{(m-1)!} \right). \end{aligned}$$

The final inequality is obtained by applying Corollary 7. The results of our experiment when using this a priori error bound are given in Figure 6.5.

As we can see, the behavior here is very similar to that observed for the matrix exponential in Figure 6.4: the a priori error estimate tends to overestimate the number of block Krylov steps required to reach the desired tolerance, whereas the simplistic error bound from Algorithm 2 follows the true error in each iteration closely.

Experiment 3 (sparse test matrices). In this experiment we compare Algorithm 2 and the block algorithm for matrices of larger scale. We use examples from the University of Florida Sparse Matrix Collection [4] and the `MUFC Twitter` matrix.[1]

From Table 6.1 we can see that the results for larger scale matrices are similar to those seen in Experiment 1. We choose *single* precision as the desired accuracy, which is achieved in all of the examples. In order to get an estimate for the error, we compute a reference solution by calling the block algorithm with *double* precision accuracy.

Some further comments are in order for the `MUFC Twitter` matrix, found in the final row of Table 6.1. This matrix is of size 148918 × 148918 with 193032 nonzero

[1]Data provided under a Creative Commons licence by The University of Strathclyde and Bloom Agency, http://www.mathstat.strath.ac.uk/outreach/twitter/mufc.

TABLE 6.1
Relative error in the 2-norm and the time (in seconds) for various test matrices when aiming for single precision accuracy.

	Alg. 2		Block alg.		
Matrix	Time	2-norm err.	Time	2-norm err.	n
Gleich/minnesota	6.00e-02	3.90e-09	2.46e+00	3.19e-10	2642
FEMLAB/poisson2D	8.32e-03	4.93e-10	5.75e-02	9.33e-11	367
Pajek/USAir97	1.32e-02	2.73e-09	5.17e-02	1.64e-09	332
Gset/G45	5.31e-02	8.46e-09	4.32e-01	2.51e-10	1000
HB/gre_1107	1.36e-02	1.24e-09	5.08e-01	7.02e-10	1107
vanHeukelum/cage8	2.20e-02	4.83e-09	3.98e-01	1.42e-10	1015
MUFC Twitter	4.39e+00	1.16e-08	6.17e+00	1.40e-08	148918

TABLE 6.2
Relative errors and condition numbers errors when computing the exponential of large sparse matrices using Algorithm 6 for the 2-norm and [5, Alg. 4.1] for the 1-norm.

Matrix	2-norm err.	2-norm cond u	1-norm err.	1-norm cond u
Gleich/minnesota	9.17e-11	1.11e-7	3.62e-11	3.18e-7
FEMLAB/poisson2D	4.34e-11	2.42e-8	4.01e-11	4.26e-8
Pajek/USAir97	9.14e-10	3.24e-7	7.68e-10	1.97e-6
Gset/G45	1.0	4.88e+4	1.0	5.48e-6
HB/gre_1107	1.46e-10	3.65e-8	1.09e-10	5.84e-8
vanHeukelum/cage8	3.29e-10	2.59e-7	1.25e-10	7.41e-7

entries, and therefore the rank-1 direction term matrix E in the Fréchet derivative (which is dense) needs approximately 165.2 GB of storage. As a result, using the `expmv` algorithm directly is not possible on a standard workstation. The result given for the block algorithm in Table 6.1 for this matrix required some modification of the `expmv` algorithm, allowing it to use only matrix-vector products.

Experiment 4 (condition numbers). Our final experiment shows how our algorithm can be used to bound the condition number of computing $\mathrm{e}^A b$ by combining Algorithms 3 and 6. For each of the matrices in Table 6.1, minus the MUFC Twitter, which was too large for the intermediate calculations during Code Fragment 9 to fit in memory, we report the relative error obtained when comparing our new algorithm to the exact solution and the estimated condition number (multiplied by the machine epsilon $u = 2^{-53}$) in Table 6.2. The exact solution is computed using variable precision arithmetic, as explained at the beginning of this section. In order to avoid overflow, the matrices were scaled such that $\|A\|_2 = 1$, and the vectors b were chosen to have elements sampled from a normal $N(0,1)$ distribution. In all cases we can see that the relative error is bounded above by the condition number times u, meaning that our method behaves in a forwards stable manner.

For comparison we have also included the 1-norm relative error and 1-norm condition number, where the latter is computed via the Python code corresponding to [5, Alg. 4.1].[2] It is interesting to note that, for the extremely ill-conditioned matrix "Gset/G45," estimating the condition number with Algorithm 6 would inform the user that the computed solution is untrustworthy, while the 1-norm algorithm returns a condition number estimate that is orders of magnitude smaller than the observed relative error. This discrepancy is likely due to the design choice mentioned by Deadman [5, p. 13] of choosing only a single set of parameters for computing all the matrix

[2]Download available at https://github.com/edvindeadman/fAbcond; accessed on 14.12.2015.

exponentials required by the algorithm: clearly this can lead to poor condition number estimates in some cases.

7. Conclusions. We have derived a new method for computing the action of the Fréchet derivative of a matrix function on a vector for low rank direction matrices E, making use of block Krylov subspace approximations. After introducing the method, we showed how one can obtain rigorous a priori error bounds to ensure that the action of the Fréchet derivative is computed to sufficient accuracy. Our numerical experiments confirm that our error bounds are effective in practice. Furthermore, we observe that typically the size of the required Krylov subspace is much smaller than the size of the original matrix. This allows the action of the Fréchet derivative to be computed much more efficiently than by other currently used methods.

We have also shown how our algorithm can be used to efficiently estimate the condition number of computing $f(A)b$ in the 2-norm. The growing use of matrix functions in a variety of areas within applied mathematics makes it vital to have efficient condition number estimates available. Our experiments also show that our condition number estimator can detect ill-conditioned problems, where the returned solution is likely to be suspect, that are not detected by competing algorithms.

Looking towards the future, the block Krylov scheme used by our algorithms is relatively basic at this stage. By incorporating data reuse, implicit restarting, rational Krylov approximations, and other state-of-the-art techniques, it is likely that the action of the Fréchet derivative could be computed even more efficiently. This would have a dramatic impact on the computation of the condition number, for which multiple Fréchet derivative computations are required.

REFERENCES

[1] A. H. Al-Mohy and N. J. Higham, *Computing the action of the matrix exponential, with an application to exponential integrators*, SIAM J. Sci. Comput., 33 (2011), pp. 488–511, https://doi.org/10.1137/100788860.

[2] M. Caliari, P. Kandolf, A. Ostermann, and S. Rainer, *Comparison of software for computing the action of the matrix exponential*, BIT, 54 (2014), pp. 113–128, https://doi.org/10.1007/s10543-013-0446-0.

[3] M. Caliari, M. Vianello, and L. Bergamaschi, *Interpolating discrete advection-diffusion propagators at Leja sequences*, J. Comput. Appl. Math., 172 (2004), pp. 79–99, https://doi.org/10.1016/j.cam.2003.11.015.

[4] T. Davis and Y. Hu, *The University of Florida Sparse Matrix Collection*, 2011, http://www.cise.ufl.edu/research/sparse/matrices/.

[5] E. Deadman, *Estimating the condition number of $f(A)b$*, Numer. Algorithms, 70 (2014), pp. 287–308, https://doi.org/10.1007/s11075-014-9947-4.

[6] E. Estrada and D. J. Higham, *Network properties revealed through matrix functions*, SIAM Rev., 52 (2010), pp. 696–714, https://doi.org/10.1137/090761070.

[7] E. Estrada, D. J. Higham, and N. Hatano, *Communicability betweenness in complex networks*, Phys. A, 388 (2009), pp. 764–774, https://doi.org/10.1016/j.physa.2008.11.011.

[8] S. Güttel, *Rational Krylov approximation of matrix functions: Numerical methods and optimal pole selection*, GAMM Mitt., 36 (2013), pp. 8–31, https://doi.org/10.1002/gamm.201310002.

[9] N. J. Higham, *The Matrix Computation Toolbox*, http://www.maths.manchester.ac.uk/~higham/mctoolbox.

[10] N. J. Higham, *Functions of Matrices: Theory and Computation*, SIAM, Philadelphia, PA, 2008, https://doi.org/10.1137/1.9780898717778.

[11] N. J. Higham and F. Tisseur, *A block algorithm for matrix 1-norm estimation, with an application to 1-norm pseudospectra*, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 1185–1201, https://doi.org/10.1137/S0895479899356080.

[12] D. Hipp, M. Hochbruck, and A. Ostermann, *An exponential integrator for non-autonomous parabolic problems*, Electron. Trans. Numer. Anal., 41 (2014), pp. 497–511.

[13] M. Hochbruck, C. Lubich, and H. Selhofer, *Exponential integrators for large systems of differential equations*, SIAM J. Sci. Comput., 19 (1998), pp. 1552–1574, https://doi.org/10.1137/S1064827595295337.

[14] M. Hochbruck and A. Ostermann, *Exponential integrators*, Acta Numer., 19 (2010), pp. 209–286, https://doi.org/10.1017/S0962492910000048.

[15] D. L. Michels, G. A. Sobottka, and A. G. Weber, *Exponential integrators for stiff elastodynamic problems*, ACM Trans. Graph., 33 (2014), pp. 7:1–7:20, https://doi.org/10.1145/2508462.

[16] A. Ruhe, *Implementation aspects of band Lanczos algorithms for computation of eigenvalues of large sparse symmetric matrices*, Math. Comp., 33 (1979), pp. 680–687, https://doi.org/10.1090/S0025-5718-1979-0521282-9.

[17] Y. Saad, *Analysis of some Krylov subspace approximations to the matrix exponential operator*, SIAM J. Numer. Anal., 29 (1992), pp. 209–228, https://doi.org/10.1137/0729014.

[18] Y. Saad, *Iterative Methods for Sparse Linear Systems*, 2nd ed., SIAM, Philadelphia, PA, 2003, https://doi.org/10.1137/1.9780898718003.