ADVANCES IN APPLIED MATHEMATICS 16, 321–375 (1995)

Derivatives of the Matrix Exponential and Their Computation

Igor Najfeld and Timothy F. Havel

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115-5718

Matrix exponentials and their derivatives play an important role in the perturbation analysis, control, and parameter estimation of linear dynamical systems. The well-known integral representation of the derivative of the matrix exponential $\exp(t\mathbf{A})$ in the direction $\mathbf{V}$, namely $\int_0^t \exp((t-\tau)\mathbf{A})\mathbf{V}\exp(\tau\mathbf{A})\,d\tau$, enables us to derive a number of new properties for it, along with spectral, series, and exact representations. Many of these results extend to arbitrary analytic functions of a matrix argument, for which we have also derived a simple relation between the gradients of their entries and the directional derivatives in the elementary directions. Based on these results, we construct and optimize two new algorithms for computing the directional derivative. We have also developed a new algorithm for computing the matrix exponential that is more efficient than direct Padé approximation, which is based on a rational representation of the exponential in terms of the hyperbolic function $\mathbf{A}\coth(\mathbf{A})$. Finally, these results are illustrated by an application to a biologically important parameter estimation problem which arises in nuclear magnetic resonance spectroscopy.

1. INTRODUCTION

The matrix exponential is one of the most widely used functions in applied mathematics. As a consequence, its computation has been the subject of numerous studies. Various derivatives of matrix exponentials also turn up in many applications, but far less is known about their properties and computation. The present study of the exponential and its derivatives was motivated by an application to structural molecular biology, in which the spatial structures of proteins and other biological macromolecules are computed from distance measurements obtained by nuclear magnetic resonance (NMR) spectroscopy. In order to obtain the best possible measurements, it is important to accurately estimate the relevant parameters from the available spectroscopic data. This in turn

0196-8858/95 $12.00

involves the computation of the exponentials of many large matrices, together with their derivatives.

1.1. *The Computation of Matrix Exponentials*

One of the earliest expositions on computing matrix exponentials appeared in the book by Frazer, Duncan, and Collar [25]. Since that time the number and variety of methods available have grown greatly, particularly in the 1960s and 1970s, driven primarily by their applications to structural, electrical, and control engineering as well as computational physics. Many of these methods were reviewed by Moler and Van Loan in 1978 [45], the two most promising of which were based on Schur's decomposition and on the "scaling and squaring" technique. The latter technique is an example of a wider class of methods consisting of argument reduction, approximation, and recovery. Because of their importance in this paper, we begin with a brief account of how these methods operate.

The most efficient method of calculating scalar exponentials decomposes the argument as $e^x = 2^{m+b}$, where m is an integer and $|b| \leq \frac{1}{2}$. The mantissa can be rapidly computed via an optimized polynomial or rational approximation to 2^b, after which the binary exponent need only be shifted by m bits. This approach can be extended to a matrix argument $\mathbf{A} \in \mathbf{C}^{n \times n}$ in two ways. First, if the matrix is scaled by a sufficiently large binary power $\mathbf{B} = \mathbf{A}/2^m$ to enable $\exp(\mathbf{B})$ to be computed from a low-order polynomial or rational approximation to it, then $\exp(\mathbf{A}) = \exp(\mathbf{B})^{2^m}$ can be recovered by iteratively squaring the matrix $\exp(\mathbf{B})$ m times. The efficiency and accuracy of this approach clearly depend upon a careful balance of the exponent m with the order of the approximation used. Second, if one can find an additive decomposition $\mathbf{A} = \mathbf{B} + \mathbf{C}$ with $[\mathbf{B}, \mathbf{C}] = \mathbf{0}$ such that $\|\mathbf{C}\|$ is small and $\mathbf{B}$ has a special structure which enables its exponential to be computed rapidly, then the exponential is given by the product $\exp(\mathbf{A}) = \exp(\mathbf{B})\exp(\mathbf{C})$. Unfortunately, such additive decompositions are not easy to find for arbitrary matrices $\mathbf{A}$, although in principle the commutativity condition can be relaxed by means of the Lie product formula, the Zassenhaus formula, or their variations [19, 42, 70].

A variety of improvements, variations, and alternative approaches have been developed since Moler and Van Loan's review, including, for example, Nuding's continued fraction approximations to $\tanh(\mathbf{A})$ [46], Parlett and Ng's block clustering of eigenvalues in the Schur decomposition [49], the Richardson extrapolation procedure in Walz [67], and the arithmetic–geometric-mean iteration of Stickel [59]. In Section 2 of this paper we describe two further advances in argument reduction and approximation. The first of these is a *spectral product representation* for the matrix

exponential, which enables us to split the argument **A** into a sum of two commuting matrices **B** + **C** such that both **C** and exp(**C**) are expressed in terms of dyadic matrices. The second improvement represents the exponential as a rational function of the hyperbolic matrix function **H**(**A**) = **A** coth(**A**), since its Taylor or Padé approximations depend on only even powers of **A** and converge more rapidly than those of either the exponential or tanh(**A**). The optimal combination of truncation order and scaling factor to use with the latter method is also given.

1.2. *The First Directional Derivative of the Matrix Exponential*

The majority of this paper is concerned with the properties, representations, and computation of various derivatives of the matrix exponential. The effect of perturbing the matrix exponential in the direction **V** is given by the limit

$$\mathbf{D}_{\mathbf{V}}(t,\mathbf{A}) = \lim_{h\to 0} \frac{1}{h}\left(e^{t(\mathbf{A}+h\mathbf{V})} - e^{t\mathbf{A}}\right). \tag{1}$$

where **A** and **V** are square matrices of the same size and h, $t \in \mathbf{R}$. Both the existence and the uniqueness of this limit follow from the integral representation (2) given below. We shall call $\mathbf{D}_{\mathbf{V}}(t,\mathbf{A})$ the *first directional derivative* (of the matrix exponential) evaluated at **A** in the direction **V**; this is also known as the first Gateaux derivative, although the terms polar [42], Fréchet [53], and parametric [70] derivative have been used. In addition, the first directional derivative of an arbitrary analytic function of a matrix argument $\mathbf{F}(\mathbf{A}) = \sum_{k=0}^{\infty} c_k \mathbf{A}^k$ will be denoted by $\mathbf{D}_{\mathbf{V}}(\mathbf{F}(\mathbf{A}))$, while the abbreviation $\mathbf{D}_{\mathbf{V}}(t,\mathbf{A})$ will be reserved for $\mathbf{D}_{t\mathbf{V}}(\exp(t\mathbf{A}))$.

A much more interesting representation of $\mathbf{D}_{\mathbf{V}}(t,\mathbf{A})$ is given by the definite integral

$$\mathbf{D}_{\mathbf{V}}(t,\mathbf{A}) = \int_0^t e^{(t-\tau)\mathbf{A}}\mathbf{V}e^{\tau\mathbf{A}}\,d\tau, \tag{2}$$

which simplifies to $\mathbf{D}_{\mathbf{V}}(t,\mathbf{A}) = t\mathbf{V}e^{t\mathbf{A}}$ when the commutator $[\mathbf{V},\mathbf{A}]$ vanishes. This representation first appeared in Karplus and Schwinger [36], and then in Feynman [22], Bellman [6], and Snider [58]. A further generalization to the perturbation of semigroups generated by operators in Banach spaces may be found in [37]. This representation can easily be derived from the fixed-point characterization of the solutions to the Volterra integral equation [6, 22, 36], as well as from the Volterra product integral [19, 26].

More generally, the integral representation of the kth directional derivative $\mathbf{D}_{\mathbf{V}}^{[k]}(t,\mathbf{A})$ in the direction $\mathbf{V}$ is given by the recurrence relation

$$\mathbf{D}_{\mathbf{V}}^{[k]}(t,\mathbf{A}) = k\int_0^t e^{(t-\tau)\mathbf{A}}\mathbf{V}\mathbf{D}_{\mathbf{V}}^{[k-1]}(\tau,\mathbf{A})\,d\tau \qquad \left(\mathbf{D}_{\mathbf{V}}^{[0]}(t,\mathbf{A}) = e^{t\mathbf{A}}\right). \quad (3)$$

These derivatives also appear in the convergent power series expansion [6]

$$e^{t(\mathbf{A}+h\mathbf{V})} = \sum_{k=0}^{\infty} \frac{h^k}{k!}\mathbf{D}_{\mathbf{V}}^{[k]}(t,\mathbf{A}). \quad (4)$$

We shall find it useful to have an expression for the first directional derivative of an integral matrix power $\mathbf{A}^p$ ($p > 0$). This *power directional derivative* is given by

$$\mathbf{D}_{\mathbf{V}}(\mathbf{A}^p) = \lim_{h\to 0}\frac{1}{h}\left((\mathbf{A}+h\mathbf{V})^p - \mathbf{A}^p\right) = \sum_{k=1}^{p}\mathbf{A}^{p-k}\mathbf{V}\mathbf{A}^{k-1}, \quad (5)$$

and reduces to $p\mathbf{V}\mathbf{A}^{p-1}$ when $[\mathbf{V},\mathbf{A}] = \mathbf{0}$. Equation (5), which has been attributed to Hausdorff [29], may be regarded as a discrete analogue of the integral representation (2). More generally, the second power directional derivative may be expressed in the form

$$\mathbf{D}_{\mathbf{V}}^{[2]}(\mathbf{A}^p) = 2\sum_{k=2}^{p}\mathbf{A}^{p-k}\mathbf{V}\mathbf{D}_{\mathbf{V}}^{[1]}(\mathbf{A}^{k-1}) = 2\sum_{k=2}^{p}\mathbf{A}^{p-k}\mathbf{V}\sum_{i=2}^{k}\mathbf{A}^{k-i}\mathbf{V}\mathbf{A}^{i-2}, \quad (6)$$

and as shown in Section 4.3 there is a recurrence relation between all consecutive derivatives.

The study of commutator powers $\{\mathbf{V},\mathbf{A}^k\} = [\ldots[[\mathbf{V},\mathbf{A}],\mathbf{A}],\ldots,\mathbf{A}]$ (k iterated commutators) and commutator series $\{\mathbf{V},\mathbf{F}(\mathbf{A})\} = \sum_{k=0}^{\infty}c_k\{\mathbf{V},\mathbf{A}^k\}$ dates back to Baker [4, 5] and Hausdorff [29]. They were subsequently used to obtain exponential representations of the solutions to ordinary linear differential equations by Magnus [42]. Some of their applications to quantum mechanics may be found in a review by Wilcox [70], and further generalizations to formal differential equations occur in the work of K. T. Chen [13, 14]. In Section 3 of this paper we pursue the close relationship between the Kronecker sum and commutator series studied by Powers [53], and derive a new convergence criterion for commutator series which involves the spread of the matrix $\mathbf{A}$, $\sigma(\mathbf{A}) = \max_{i,j}|\lambda_i - \lambda_j|$, where λ_i denotes its eigenvalues.

The eigenvalues of $\mathbf{D}_{\mathbf{V}}(t,\mathbf{A})$ were first computed by Magnus [42], and some early references to the eigenvalues of $\mathbf{D}_{\mathbf{V}}(\mathbf{F}(\mathbf{A}))$ may be found in Powers [52] and Flanders [24]. Since then several attempts have been made

to obtain a full spectral representation for $\mathbf{D}_{\mathbf{V}}(t,\mathbf{A})$, for instance, by Atherton and De Gance [3] and by Deif [18]. Assuming that $\mathbf{A}$ has a full set of linearly independent eigenvectors, several compact spectral representations of $\mathbf{D}_{\mathbf{V}}(t,\mathbf{A})$ in terms of the eigensystem of $\mathbf{A}$ and the Hadamard product are presented in Section 4 of this paper. These results are then generalized to the directional derivative of any analytic function as

$$\mathbf{D}_{\mathbf{V}}(\mathbf{F}(\mathbf{A})) = \mathbf{U}\big(\bar{\mathbf{V}} \odot \Delta_{\lambda}(\mathbf{F})\big)\mathbf{U}^{-1}. \tag{7}$$

where $\bar{\mathbf{V}} = \mathbf{U}^{-1}\mathbf{V}\mathbf{U}$, and $\Delta_{\lambda}(\mathbf{F})$ is the Loewner matrix of $\mathbf{F}$ evaluated at the eigenvalues of $\mathbf{A}$ (see Theorem 4.8).

The representation of $\mathbf{D}_{\mathbf{V}}(t,\mathbf{A})$ in terms of the product of $\exp(t\mathbf{A})$ with the commutator series $\{\mathbf{V},(\exp(t\mathbf{A})-\mathbf{I})/\mathbf{A}\}$ first appeared in Hausdorff [29] and was later expressed in terms of Kronecker sums and generalized to arbitrary analytic functions by Powers [53]. His representation of the first directional derivative in terms of the commutator powers is given by

$$\mathbf{D}_{\mathbf{V}}(\mathbf{F}(\mathbf{A})) = \sum_{k=1}^{\infty} \frac{\mathbf{F}^{[k]}(\mathbf{A})}{k!}\{\mathbf{V},\mathbf{A}^{k-1}\}, \tag{8}$$

where $\mathbf{F}^{[k]}(\mathbf{A})$ is the kth formal derivative of $\mathbf{F}$ evaluated at $\mathbf{A}$. In the case of the matrix exponential all the formal derivatives are the same and Eq. (8) reduces to the Hausdorff series. In Section 4 of this paper, the Hausdorff commutator series is summed to obtain a new closed form representation

$$\operatorname{col}(\mathbf{D}_{\mathbf{V}}(t,\mathbf{A})) = (\mathbf{I}\otimes e^{t\mathbf{A}})\left(\frac{e^{t\mathbf{K}(\mathbf{A})}-\mathbf{I}}{\mathbf{K}(\mathbf{A})}\right)\operatorname{col}(\mathbf{V}), \tag{9}$$

where $\mathbf{K}(\mathbf{A}) = \mathbf{A}^{\mathrm{T}} \oplus (-\mathbf{A})$ and col stacks the columns of a matrix.

In Section 4 we also derive another new representation of $\mathbf{D}_{\mathbf{V}}(t,\mathbf{A})$ in terms of a faster converging "sinch" commutator series $\{\mathbf{V},\sinh(t\mathbf{A})/t\mathbf{A}\}$. This is used in Section 5 to obtain a new "scaling and squaring" algorithm for $\mathbf{D}_{\mathbf{V}}(t,\mathbf{A})$, by evaluating the truncated series at the scaled argument $\mathbf{B} = t\mathbf{A}/2^m$, and then using the "two-sided" doubling relation $\mathbf{D}_{\mathbf{V}}(2,\mathbf{B}) = \exp(\mathbf{B})\mathbf{D}_{\mathbf{V}}(1,\mathbf{B}) + \mathbf{D}_{\mathbf{V}}(1,\mathbf{B})\exp(\mathbf{B})$ (see, e.g., Ref. [38]) to recover the derivative at $t\mathbf{A}$. The optimal combination of exponent m and truncation order for this algorithm is also determined.

Because an analytic function of a $k \times k$ Jordan block is a triangular as well as a semi-circulant matrix, it explicitly displays the values of the function and its first $k-1$ derivatives along the first row. It is therefore not surprising that functions of other special triangular matrices will lead to other derivatives and their discrete analogues, divided differences.

Examples may be found in the results of Rosenbloom [57], Davis [16], and Opitz [47] as well as in the recurrence relations for analytic functions of triangular matrices given in Parlett [48], Filliponi [23], and Van Loan [62]. Another interesting example may be found in Jernigan's work on polymer statistical mechanics [35]. We shall show in Section 4 of this paper that an analytic function **F** evaluated at the block triangular matrix **B**

$$\mathbf{F}(\mathbf{B}) = \mathbf{F}\left(\begin{bmatrix} \mathbf{A} & \mathbf{V} \\ \mathbf{0} & \mathbf{A} \end{bmatrix}\right) = \begin{bmatrix} \mathbf{F}(\mathbf{A}) & \mathbf{D}_{\mathbf{V}}(\mathbf{F}(\mathbf{A})) \\ \mathbf{0} & \mathbf{F}(\mathbf{A}) \end{bmatrix} \tag{10}$$

explicitly contains its first directional derivative in the above-diagonal block. This result is then generalized to the first q directional derivatives of **F** by means of the block bidiagonal semi-circulant matrix **B**

$$\mathbf{F}(\mathbf{B}) = \mathbf{F}\left(\begin{bmatrix} \mathbf{A} & \mathbf{V} & & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{A} & \mathbf{V} & \cdots & \mathbf{0} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{0} & & & \cdots & \mathbf{A} \end{bmatrix}\right)$$

$$= \begin{bmatrix} \mathbf{F}(\mathbf{A}) & \dfrac{\mathbf{D}_{\mathbf{V}}^{[1]}(\mathbf{F}(\mathbf{A}))}{1!} & \cdots & \dfrac{\mathbf{D}_{\mathbf{V}}^{[q]}(\mathbf{F}(\mathbf{A}))}{q!} \\ \mathbf{0} & \mathbf{F}(\mathbf{A}) & \cdots & \dfrac{\mathbf{D}_{\mathbf{V}}^{[q-1]}(\mathbf{F}(\mathbf{A}))}{(q-1)!} \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{F}(\mathbf{A}) \end{bmatrix}, \tag{11}$$

where the number of **V** blocks along the first superdiagonal is equal to q, the highest-order derivative desired. We call **B** the qth order *directional Jordan block* defined by **A** and **V**. In Section 5, Eq. (10) will be used to obtain a second "scaling and squaring" algorithm for the first directional derivative of the matrix exponential.

The gradients of the entries of a matrix exponential with respect to its argument, $d\exp_{ij}(\mathbf{A})/d\mathbf{A}$, have been studied by Athans and Schweppe [2], Brewer [8], Graham [27], and Vetter [64]. In Section 6, we derive a relation between the gradients of the entries of an analytic matrix function **F**(**A**) and its directional derivatives in the elementary directions $\mathbf{E}_{ij}$ (with a one in the ijth entry and zeros elsewhere). This relation is specialized to the case of symmetric matrices.

Directional derivatives can also be represented by a generalization of the Cauchy integral formula known as the Dunford–Taylor integral. Although we shall not pursue this approach here, we note early results by Reinhart [54, 55] and perturbations of resolvents by Kato [37]. These

results were later extended by Stickel [60] to the following representation of the directional derivative of an analytic matrix valued function,

$$\mathbf{D}_{\mathbf{V}}(\mathbf{F}(\mathbf{A})) = \frac{1}{2\pi i}\int_{\Gamma} f(z)(z\mathbf{I} - \mathbf{A})^{-1}\mathbf{V}(z\mathbf{I} - \mathbf{A})^{-1}\,dz. \tag{12}$$

where f is the scalar version of $\mathbf{F}$, while Γ is a closed curve enclosing all the eigenvalues of $\mathbf{A}$ and whose winding number with respect to each eigenvalue is equal to one.

1.3. *Selected Applications of Matrix Exponential Derivatives*

One of the most frequent applications of the integral representation (2) has been in quantum mechanics, where it appears to have been discovered in a slightly more general form (see, e.g., reviews by Wilcox [70] and Lutzky [41]). Let the operator $\mathbf{Z}(\zeta)$ be an analytic function of a parameter $\zeta \in \mathbf{D} \subseteq \mathbf{C}$ such that both $d\mathbf{Z}/d\zeta = \mathbf{Z}'(\zeta)$ and $\exp(\mathbf{Z}(\zeta))$ are defined over the open set $\mathbf{D}$. The operator $d(\exp(\mathbf{Z}(\zeta)))/d\zeta$ is called the *parametric derivative*, and the integral representation of the first parametric derivative at ζ is given by

$$\frac{de^{\mathbf{Z}(\zeta)}}{d\zeta} = \int_0^1 e^{(1-\tau)\mathbf{Z}(\zeta)}\mathbf{Z}'(\zeta)e^{\tau\mathbf{Z}(\zeta)}\,d\tau. \tag{13}$$

Thus the parametric derivative in Eq. (13) is equal to the directional derivative $\mathbf{D}_{\mathbf{Z}'(\zeta)}(1, \mathbf{Z}(\zeta))$, where $\mathbf{Z}'(\zeta)$ can be viewed as a direction of perturbation in the expansion $\mathbf{Z}(\zeta + h) = \mathbf{Z}(\zeta) + h\mathbf{Z}'(\zeta) + \cdots$.

The applications of matrix calculus to the optimization and the sensitivity analysis of linear control systems are explored by Athans and Schweppe [2], Atherton and De Gance [3], Vetter [63–65], and Brewer [8–12]. In particular, Brewer [9] has given a generalization of Eq. (13) to the derivative of $\exp(\mathbf{W}(\mathbf{Q}))$ with respect to a matrix argument $\mathbf{Q}$

$$\frac{\partial e^{t\mathbf{W}(\mathbf{Q})}}{\partial \mathbf{Q}} = \int_0^t \left(\mathbf{I}_p \otimes e^{(t-\tau)\mathbf{W}(\mathbf{Q})}\right)\frac{\partial \mathbf{W}}{\partial \mathbf{Q}}\left(\mathbf{I}_q \otimes e^{\tau\mathbf{W}(\mathbf{Q})}\right)d\tau, \tag{14}$$

where $\mathbf{W}(\mathbf{Q})$ is an $n \times n$ matrix that depends on a $p \times q$ matrix $\mathbf{Q}$, the derivative matrices $\partial\mathbf{W}/\partial\mathbf{Q}$ and $\partial \exp(t\mathbf{W}(\mathbf{Q}))/\partial\mathbf{Q}$ have size $np \times nq$, $\mathbf{I}_p$ is an identity matrix of size p, and $\otimes$ denotes the Kronecker product.

The sensitivity of the matrix exponential to perturbations in its matrix argument provides an estimate of the accuracy with which it can be computed. This is usually quantitated by means of the exponential condi-

tion number defined in Van Loan [61], which is given by

$$\nu(t, \mathbf{A}) = \frac{\|\mathbf{D}(t, \mathbf{A})\| \, \|\mathbf{A}\|}{\|e^{t\mathbf{A}}\|}, \tag{15}$$

where $\mathbf{D}(t, \mathbf{A}) = \sup_{\|\mathbf{V}\| \leq 1}(\mathbf{D}_{\mathbf{V}}(t, \mathbf{A}))$. This was further studied by Roche [56], who found that the sensitivity depends on the clustering of the eigenvalues as well as non-normality. Spectral methods of estimating this condition number have been given by Kenney and Laub [38], while estimates obtained by applying standard numerical integration techniques to the integral representation (2) have been given by Kenney and Laub [38] as well as Mathias [44].

Parameter estimation in time-invariant linear systems has recently proved important in several biological applications. In the least-squares approach, one minimizes the sum of the squares χ^2 of the differences between the calculated and observed values of the solution to the equations at a number of fixed time points. For example, Jennrich and Bright [34] have used this approach to estimate the rates of diffusion of a biologically active molecule through the "compartments" of an organism's body. In another example, Koehl and Lefèvre [39] have estimated "relaxation rates" from NMR data, while Yip and Case [72] have used similar methods to refine molecular structures so that the parameters computed from their Cartesian coordinates minimize the same sum of squares.

Efficient methods of minimizing χ^2 require at least its gradient $\nabla\chi^2$ with respect to the parameters, which in turn leads to the problem of evaluating the gradients $d\exp_{ij}(\mathbf{A})/d\mathbf{A}$ of the entries of the matrix exponential. Koehl and Lefèvre have used a truncated Hausdorff expansion (see Section 4.3) to evaluate the latter gradients, while spectral representations have been derived by Jennrich and Bright as well as by Yip and Case. More recently, Yip has also used numerical integration of the integral representation [73]. In Section 7, we show how the results obtained in this paper enable one to compute the total gradient $\nabla\chi^2$ yet more efficiently.

2. THE COMPUTATION OF THE MATRIX EXPONENTIAL

Argument reduction, approximation, and recovery (see Section 1.1) constitute one of the most general and reliable approaches to computing the matrix exponential $\exp(\mathbf{A})$. By far the most widely used version of this approach is the "scaling and squaring" method. The number of matrix multiplications required by this method, however, grows as $\log\|\mathbf{A}\|$, and problems with round-off error propagation can occur in the process. This

part of the paper presents several improvements on the overall reduction, approximation, and recovery approach that reduce the total amount of work required, and usually increase the accuracy of the result as well.

Although argument reduction by scaling can always be used, additive decompositions are potentially more effective. The idea is to decompose the argument as $\mathbf{A} = \mathbf{B} + \mathbf{C}$, where $\mathbf{B}$ is a relatively large matrix whose structure is simple so that $\exp(\mathbf{B})$ can be evaluated very rapidly, and $\|\mathbf{C}\|$ is small enough to enable $\exp(\mathbf{C})$ to be approximated without much computational effort. The major obstacle is that $[\mathbf{B}, \mathbf{C}] \neq \mathbf{0}$ for most decompositions, so that $\exp(\mathbf{A}) \neq \exp(\mathbf{B})\exp(\mathbf{C})$. It follows that we must look for a matrix $\mathbf{B}$ in the *centralizer* of $\mathbf{A}$, that is, the subalgebra of all matrices that commute with $\mathbf{A}$. Although the explicit construction of the centralizer would require the complete diagonalization of $\mathbf{A}$, in the case of semi-simple matrices the centralizer consists of all polynomials in $\mathbf{A}$ of degree at most $n - 1$ [17]. The problem with constructing $\mathbf{B}$ as a polynomial in $\mathbf{A}$ is that it is usually either not particularly simple or else $\|\mathbf{A} - \mathbf{B}\|$ is not sufficiently small. For example, one could use the first coefficient of the characteristic polynomial of $\mathbf{A}$, i.e., $\mathbf{B} = \operatorname{tr}(\mathbf{A})\mathbf{I}/n$, as was done in [68] with only modest gains in the general case. The reason is that the trace, being the sum of all eigenvalues, contains little information about their distribution, so that one can expect a significant reduction in $\|\mathbf{A} - \mathbf{B}\|$ only for matrices having clustered eigenvalues.

2.1. *The Spectral Product Representation*

We now derive a *spectral product* representation for the matrix exponential. Surprisingly, we have not been able to find any references to such a representation except for a very special case mentioned in [21], where it was used to derive a well-known spectral sum formula for a group of symmetric positive-definite operators. We start by writing the Jordan canonical form of $\mathbf{A}$ in several ways:

$$\mathbf{A} = \mathbf{U}\mathbf{J}\mathbf{U}^{-1} = \sum_{k=1}^{m} \mathbf{C}_k = \sum_{k=1}^{m} \mathbf{U}_k \mathbf{J}_k \mathbf{V}_k^{\mathrm{H}}$$

$$= \sum_{k=1}^{m} \mathbf{U}_k \left(\lambda_k \mathbf{I}_{n_k} + \mathbf{N}_k \right) \mathbf{V}_k^{\mathrm{H}}. \tag{16}$$

Here, $m \leq n$ is the number of Jordan blocks, n_k is the dimension of the Jordan block corresponding to the kth eigenvalue λ_k, and $\sum_{k=1}^{m} n_k = n$. The n_k columns of the matrices $\mathbf{U}_k$ and $\mathbf{V}_k$ are the left and right eigenvectors of λ_k, respectively. The matrix $\mathbf{N}_k$ is the usual $n_k \times n_k$ elementary nilpotent matrix with ones down the first upper-diagonal.

THEOREM 2.1. *If* $\mathbf{A}$ *is an arbitrary square matrix with* m *Jordan blocks, then its matrix exponential can be factored as a product of* m *commuting matrices*

$$e^{t\mathbf{A}} = \prod_{k=1}^{m} \Omega_k(t) = \prod_{k=1}^{m} \left(\mathbf{I} + \mathbf{U}_k\left(e^{t\lambda_k} e^{t\mathbf{N}_k} - \mathbf{I}_{n_k}\right)\mathbf{V}_k^{\mathsf{H}}\right), \tag{17}$$

where $\exp(t\mathbf{N}_k)$ *is given as usual* [26, 50] *by*

$$e^{t\mathbf{N}_k} = \begin{bmatrix} 1 & t & \cdots & t^{n_k-1}/(n_k-1)! \\ 0 & 1 & \cdots & t^{n_k-2}/(n_k-2)! \\ & & \cdots & \\ 0 & 0 & \cdots & 1 \end{bmatrix}. \tag{18}$$

Proof. By the biorthogonality of the left and right eigenvectors, $\mathbf{V}_k^{\mathsf{H}}\mathbf{U}_l = \mathbf{0}$ whenever $k \neq l$. It follows that $\mathbf{C}_k\mathbf{C}_l = \mathbf{0}$ for all $k \neq l$, so that $[\mathbf{C}_k, \mathbf{C}_l] = \mathbf{0}$. Therefore the exponential can be factored as

$$e^{t\mathbf{A}} = \exp\left(t \sum_{k=1}^{m} \mathbf{C}_k\right) = \prod_{k=1}^{m} e^{t\mathbf{C}_k} = \prod_{k=1}^{m} \Omega_k(t). \tag{19}$$

The expression for $\Omega_k(t)$ can be further simplified using the biorthonormality condition $\mathbf{V}_k^{\mathsf{H}}\mathbf{U}_k = \mathbf{I}_{n_k}$, as

$$\begin{aligned} \Omega_k(t) &= e^{t\mathbf{U}_k\mathbf{J}_k\mathbf{V}_k^{\mathsf{H}}} = \mathbf{I} + t\mathbf{U}_k\mathbf{J}_k\mathbf{V}_k^{\mathsf{H}} + \frac{t^2}{2!}\mathbf{U}_k\mathbf{J}_k\mathbf{V}_k^{\mathsf{H}}\mathbf{U}_k\mathbf{J}_k\mathbf{V}_k^{\mathsf{H}} + \cdots \\ &= \mathbf{I} + \mathbf{U}_k\left(e^{t\mathbf{J}_k} - \mathbf{I}_{n_k}\right)\mathbf{V}_k^{\mathsf{H}} = \mathbf{I} + \mathbf{U}_k\left(e^{t\lambda_k}e^{t\mathbf{N}_k} - \mathbf{I}_{n_k}\right)\mathbf{V}_k^{\mathsf{H}}. \quad \blacksquare \end{aligned} \tag{20}$$

COROLLARY 2.2. *If* $\mathbf{A}$ *is a semi-simple matrix (i.e.,* $m = n$ *above) and* $\lambda_k, \mathbf{u}_k, \mathbf{v}_k$ *are its eigenvalues and left and right eigenvectors, respectively, then each matrix* $\Omega_k(t)$ *in Eq.* (17) *further simplifies to*

$$\Omega_k(t) = \mathbf{I} + (e^{t\lambda_k} - 1)\mathbf{u}_k\mathbf{v}_k^{\mathsf{H}}. \tag{21}$$

Now suppose that $\mathbf{A}$ is a semi-simple matrix with a known eigenvalue λ and associated left and right eigenvectors $\mathbf{u}$ and $\mathbf{v}$. The matrix $\mathbf{A}[\lambda] = \mathbf{A} - \lambda\mathbf{u}\mathbf{v}^{\mathsf{H}} = \mathbf{A} - \mathbf{C}_\lambda$ is defined to be its *spectral reduction* (with respect to λ). The eigenstructure of $\mathbf{A}[\lambda]$ remains the same as that of $\mathbf{A}$ except that the eigenvalue λ has been replaced by zero. The k-fold spectral reduction of $\mathbf{A}$ is then defined for $k < n$ by

$$\mathbf{A}[\lambda_1, \lambda_2, \ldots, \lambda_k] = \mathbf{A} - \sum_{i=1}^{k} \mathbf{C}_i = \mathbf{A} - \sum_{i=1}^{k} \lambda_i\mathbf{u}_i\mathbf{v}_i^{\mathsf{H}}. \tag{22}$$

The matrices $\mathbf{C}_i$ $(i = 1, \ldots, k)$ commute with $\mathbf{A}[\lambda_1, \lambda_2, \ldots, \lambda_k]$ and hence determine a *partial spectral product* representation of the matrix exponential

$$e^{t\mathbf{A}} = e^{t\mathbf{A}[\lambda_1, \lambda_2, \ldots, \lambda_k]} \prod_{i=1}^{k} e^{t\mathbf{C}_i}. \tag{23}$$

If these eigenvalues satisfy $\rho(\mathbf{A}) = |\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_k|$, where $\rho(\mathbf{A})$ is the spectral radius of $\mathbf{A}$, then the following sequence of inequalities holds:

$$\rho(\mathbf{A}[\lambda_1, \lambda_2, \ldots, \lambda_k]) \leq \rho(\mathbf{A}[\lambda_1, \lambda_2, \ldots, \lambda_{k-1}]) \leq \cdots \leq \rho(\mathbf{A}). \tag{24}$$

The bound $\rho(\mathbf{A}[\lambda_1, \lambda_2, \ldots, \lambda_k]) \leq |\lambda_k|$ can be used to decide when it would be more efficient to switch from spectral product reduction to the computation of $\exp(t\mathbf{A}[\lambda_1, \lambda_2, \ldots, \lambda_k])$ by some other method.

The advantage of the partial spectral product representation lies in the fact that the first $k \ll n$ eigenvalues and eigenvectors can generally be calculated very quickly. Note also that the products of the form $\exp(t\mathbf{C}_i) \cdot \mathbf{M}$ in Eq. (23) can be evaluated in time $O(n^2)$, and that if $\mathbf{A}$ is a skew-Hermitian matrix then each of the factors $\exp(t\mathbf{C}_i)$ is unitary. As illustrated in Appendix A1, whenever it is possible to accurately calculate the largest few eigenvalues and the corresponding eigenvectors of the matrix $\mathbf{A}$, the representation (23) also has quite a beneficial effect on round-off error propagation in the computation of $\exp(t\mathbf{A})$.

2.2. *A General Algorithm for the Matrix Exponential*

When the norm of the argument matrix $\|\mathbf{A}\|$ (or its spectral reduction $\|\mathbf{A}[\lambda_1, \lambda_2, \ldots, \lambda_k]\|$) is too large for direct approximation, it must be reduced by scaling. A power of two is the simplest and most stable scale factor, although arbitrary natural numbers have been used [51]. While most approximations to the exponential of the scaled argument are based on a truncated Taylor, Padé, or continued fraction expansion, we have found it more efficient to approximate the matrix function

$$\mathbf{H}(\mathbf{A}) = \mathbf{A}\coth(\mathbf{A}) = \mathbf{A}\frac{e^{2\mathbf{A}} + \mathbf{I}}{e^{2\mathbf{A}} - \mathbf{I}}, \tag{25}$$

from which the exponential can be obtained as

$$e^{2\mathbf{A}} = \frac{\mathbf{H}(\mathbf{A}) + \mathbf{A}}{\mathbf{H}(\mathbf{A}) - \mathbf{A}}. \tag{26}$$

Here and in what follows, we shall write matrix division explicitly whenever the numerator and denominator matrices commute and the denominator

matrix is either nonsingular or the singularities in question are removable (as in Eq. (25)).

The general algorithm we advocate for computing $\exp(\mathbf{A})$ is a four stage process consisting of:

1. Calculation of an integer $d \geq 0$ such that the scaled argument $\mathbf{B} = \mathbf{A}/2^{d+1}$ satisfies $\|\mathbf{B}^2\| \leq \gamma$, where γ is a constant obtained from the optimization process in Section 2.3.

2. Rational approximation of $\mathbf{H}(\mathbf{B})$, which can be done in a fixed number of matrix multiplications and divisions.

3. Evaluation of $\exp(2\mathbf{B})$ by means of Eq. (26).

4. Iterative squaring of the result $\exp(2\mathbf{B})$ d times to recover the exponential of the original matrix $\mathbf{A}$.

Since $\mathbf{H}(\mathbf{B})$ is a matrix function that depends on only even powers of $\mathbf{B}$ (see Eq. (29) below) and whose error norm estimates can be expressed in terms of $\|\mathbf{B}^2\|$, the integer d can be computed from the inequality

$$\|\mathbf{B}^2\| = \frac{\|\mathbf{A}^2\|}{2^{2d+2}} \leq \gamma. \tag{27}$$

The real number $\gamma > 0$ depends on the chosen approximation, and must be small enough to ensure that $\mathbf{H}(\mathbf{B})$ will be approximated to the desired accuracy. Thus once γ is known, d is given by

$$d = \begin{cases} \left\lceil \frac{1}{2}\log_2(\|\mathbf{A}^2\|/4\gamma) \right\rceil & \text{if } \|\mathbf{A}^2\| > 4\gamma \\ 0 & \text{otherwise.} \end{cases} \tag{28}$$

The advantage of using $\|\mathbf{A}^2\|$ is that it results in a smaller number of iterative squarings d, particularly when $\mathbf{A}$ is nonnormal with $\|\mathbf{A}^2\| \ll \|\mathbf{A}\|^2$.

The power series for $\mathbf{H}(\mathbf{B})$ is given by

$$\mathbf{H}(\mathbf{B}) = \mathbf{I} + \frac{1}{3}\mathbf{B}^2 - \frac{1}{45}\mathbf{B}^4 + \cdots + \frac{2^{2k}B_{2k}}{(2k)!}\mathbf{B}^{2k} + \cdots, \tag{29}$$

where the B_{2k} are Bernoulli numbers. This shows that the singularities of $\mathbf{H}(\mathbf{B})$ (defined as in Eq. (25)) due to any zero eigenvalues in $\mathbf{B}$ are removable. Although this series converges absolutely only if the spectral radius of $\mathbf{B}$ satisfies $\rho(\mathbf{B}) < \pi$, Eq. (25) constitutes an analytic continuation provided that no eigenvalue of $\mathbf{B}$ is equal to an integer multiple of the purely imaginary number $i\pi$ (which are generally the nonremovable singularities in $\mathbf{H}(\mathbf{B})$). The same constraint applies to the functional equation for argument doubling $\mathbf{H}(2\mathbf{B}) = \mathbf{H}(\mathbf{B}) + \mathbf{B}^2/\mathbf{H}(\mathbf{B})$, which provides a recursive means of performing analytic continuation.

The well-known asymptotic behavior of the coefficients in the series (29), i.e.,

$$\frac{2}{(2\pi)^{2k}} < \frac{(-1)^{k+1}B_{2k}}{(2k)!} < \frac{2}{(2\pi)^{2k}}\left(1 + \frac{1}{2^{2k-1}-1}\right) \qquad (k = 1,2,\ldots) \tag{30}$$

(see, e.g., Ref. [1]), suggests that a rational or continued fraction approximation to $\mathbf{H(B)}$ should converge more rapidly than a polynomial one. The continued fraction expansion of $\mathbf{H(B)}$ is very simple, namely

$$\mathbf{H(B)} = \mathbf{I} + \cfrac{\mathbf{B}^2}{3\mathbf{I} + \cfrac{\mathbf{B}^2}{5\mathbf{I} + \cfrac{\mathbf{B}^2}{7\mathbf{I} + \cdots}}}, \tag{31}$$

and may be used to derive the rational approximations in the diagonal and superdiagonal entries of its Padé table. The diagonal approximations are particularly efficient due to the equal numbers of powers in the numerator and denominator polynomials. The first four of these rational approximations were computed using the MAPLE symbolic algebra program, and are given below:

$$\begin{aligned}
\mathbf{H}_2(\mathbf{B}) &= \frac{\mathbf{I} + \frac{2}{5}\mathbf{B}^2}{\mathbf{I} + \frac{1}{15}\mathbf{B}^2} \\
\mathbf{H}_4(\mathbf{B}) &= \frac{\mathbf{I} + \frac{4}{9}\mathbf{B}^2 + \frac{1}{63}\mathbf{B}^4}{\mathbf{I} + \frac{1}{9}\mathbf{B}^2 + \frac{1}{945}\mathbf{B}^4} \\
\mathbf{H}_6(\mathbf{B}) &= \frac{\mathbf{I} + \frac{6}{13}\mathbf{B}^2 + \frac{10}{429}\mathbf{B}^4 + \frac{4}{19305}\mathbf{B}^6}{\mathbf{I} + \frac{5}{39}\mathbf{B}^2 + \frac{2}{715}\mathbf{B}^4 + \frac{1}{135135}\mathbf{B}^6} \\
\mathbf{H}_8(\mathbf{B}) &= \frac{\mathbf{I} + \frac{8}{17}\mathbf{B}^2 + \frac{7}{255}\mathbf{B}^4 + \frac{4}{9945}\mathbf{B}^6 + \frac{1}{765765}\mathbf{B}^8}{\mathbf{I} + \frac{7}{51}\mathbf{B}^2 + \frac{1}{255}\mathbf{B}^4 + \frac{2}{69615}\mathbf{B}^6 + \frac{1}{34459425}\mathbf{B}^8}.
\end{aligned} \tag{32}$$

We note that if $\mathbf{A}$ is skew-Hermitian, then $\mathbf{H(A)}$ is Hermitian and it is easy to show that the right-hand side of Eq. (26) is unitary. Hence the rational representation (26) can be regarded as a generalized Cayley transform of $\mathbf{A}$. The unitary character of this transform is preserved even when $\mathbf{H(A)}$ is replaced by any polynomial or rational approximation to it. In contrast, if

one uses a direct Padé approximation to exp(2**B**), only the diagonal entries in the Padé table have this property.

Let **M** = **M**(**B**) and **N** = **N**(**B**) be the numerator and denominator polynomials in a rational approximation to **H**(**B**). The third step of the above algorithm computes

$$e^{2\mathbf{B}} = \frac{\mathbf{M}/\mathbf{N} + \mathbf{B}}{\mathbf{M}/\mathbf{N} - \mathbf{B}} = \frac{\mathbf{M} + \mathbf{BN}}{\mathbf{M} - \mathbf{BN}}. \tag{33}$$

Although there is no algebraic difference between these two expressions, there is a small difference in computational complexity which consists of the amount by which a matrix division exceeds a multiplication (about $O(2n^3/3)$ flops). Both expressions are numerically stable provided $\|\mathbf{B}^2\|$ is sufficiently smaller than π^2, which holds for the optimal approximations described in Section 2.3. Nevertheless, the last matrix division in either expression requires some care, and one step iterative refinement is recommended if matrix **B** deviates significantly from a normal one.

With the approximation to exp(2**B**) computed, the fourth step of the algorithm is iterative squaring, which is one of the simplest nonlinear recurrences, namely

$$\begin{aligned} \mathbf{E}_0 &:= e^{2\mathbf{B}} \\ \mathbf{E}_k &:= \mathbf{E}_{k-1}^2 \qquad (k = 1, \ldots, d), \end{aligned} \tag{34}$$

where d is obtained from Eq. (28). At the end of this loop, $\mathbf{E}_d$ is equal to exp(**A**) up to rounding error. The loop requires d matrix multiplications, and since d becomes large when $\frac{1}{2}\log_2(\|\mathbf{A}^2\|/4\gamma) \gg 1$, iterative squaring can be the most time consuming step of the overall computation. In such cases, the partial spectral product decomposition described above may prove useful.

Iterative squaring also tends to amplify the rounding errors, so that even when the computation of exp(2**B**) is carried out to machine precision significant errors in exp(**A**) may occur. The problem is particularly significant when **A** is a strictly stable but not a normal matrix, since then some of the entries in $\exp(2^k\mathbf{B})$ may undergo considerable growth for some values of k before they ultimately approach zero, leaving predominantly rounding errors behind. This is known as overshoot, large transients, or "the hump") [15, 45], and is in fact closely related to the Kreiss matrix theorem which bounds the powers $(\exp(2\mathbf{B}))^{2^k}$ of the exponential [69].

2.3. *The Optimal Computation of* $e^{t\mathbf{A}}$

To find the interval γ in Eq. (27) for the diagonal Padé approximates $\mathbf{H}_{2m}$ (see Eq. (32)), we first obtain a bound on the norm of the remainder

series

$$\mathbf{H}(\mathbf{B}) - \mathbf{H}_{2m}(\mathbf{B}) = \frac{\mathbf{B}^{4m+2}}{c_{2m}}\left(\mathbf{I} - b_{2m}^{[1]}\mathbf{B}^2 + b_{2m}^{[2]}\mathbf{B}^4 - b_{2m}^{[3]}\mathbf{B}^6 \cdots\right). \quad (35)$$

In bounding the norm of the right-hand side of (35), we shall treat $\gamma = \|\mathbf{B}^2\| \geq 0$ as an independent variable. This gives rise to the inequality

$$\|\mathbf{H}(\mathbf{B}) - \mathbf{H}_{2m}(\mathbf{B})\| \leq \frac{\gamma^{2m+1}}{c_{2m}} w_{2m}(\gamma), \quad (36)$$

where the coefficient c_{2m} increases with m, and all the coefficients $b_{2m}^{[j]}$ in the series $w_{2m}(\gamma) = 1 + b_{2m}^{[1]}\gamma + b_{2m}^{[2]}\gamma^2 + b_{2m}^{[3]}\gamma^3 \cdots$ are easily computed using the MAPLE program. These are found to be positive for all values of $m \leq 5$ and $j < 92$, which, as shown below, include all the values we need in order to compute the optimal approximations for single and double precision accuracy.

Since γ and the $b_{2m}^{[j]}$ are positive, $w_{2m}(\gamma)$ is a convex, monotonically increasing analytic function such that $w_{2m}(\gamma) \geq 1$ (the first singularity is approximately at π). Given any $\varepsilon > 0$, we determine the interval γ by setting the upper bound in Eq. (36) equal to it, i.e.,

$$\frac{\gamma^{2m+1}}{c_{2m}} w_{2m}(\gamma) = \varepsilon. \quad (37)$$

For each fixed m, the properties of $w_{2m}(\gamma)$ guarantee that the nonlinear equation (37) possesses a unique positive solution γ_{2m}. Furthermore, Newton's method applied to Eq. (37) generates a monotonically decreasing sequence of iterates $\gamma_{2m}^{[i]}$ that converges quadratically to this solution provided that it starts from the overestimate $\gamma_{2m}^{[0]} = (\varepsilon c_{2m})^{1/(2m+1)}$. Table I contains the numerical solutions of Eq. (37) for all consecutive values $1 \leq m \leq 5$ and the usual single and double machine precisions.

TABLE I

m	$\varepsilon = 2^{-23}$	$\varepsilon = 2^{-52}$
1	5.70678×10^{-2}	7.04539×10^{-5}
2	9.82521×10^{-1}	1.85078×10^{-2}
3	3.72354×10^{0}	2.47517×10^{-1}
4		1.15192×10^{0}
5		3.17794×10^{0}

The criterion for choosing a particular diagonal Padé approximation $\mathbf{H}_{2m}$ is the minimization of the work $\eta = \eta(m)$ required for steps 2, 3, and 4 of the algorithm in Section 2.2, in units of matrix multiplications. Assuming that these steps involve only one matrix division (see Eq. (33)), which takes work approximately equal to five-thirds matrix multiplications, we can write this as

$$\eta(m) = \tfrac{8}{3} + m + \lceil d(\gamma_{2m}) \rceil = \tfrac{5}{3} + \lceil m - \tfrac{1}{2}\log_2(\gamma_{2m}) + \tfrac{1}{2}\log_2(\|\mathbf{A}^2\|) \rceil. \tag{38}$$

since $d(\gamma_{2m}) = \frac{1}{2}\log_2(\|\mathbf{A}^2\|/4\gamma_{2m})$. The minimization of $\eta(m)$ takes place over the positive integers, and it is clear from Eq. (38) that the optimal m does not depend on $\|\mathbf{A}^2\|$. On substituting the values of γ_{2m} from Table I into this equation, it is found that the associated minima are attained at $m^* = 2$ for $\varepsilon = 2^{-23}$ and at $m^* = 4$ for $\varepsilon = 2^{-52}$. Thus the optimal single and double precision approximations are $\mathbf{H}_4(\mathbf{B})$ with $\gamma = \gamma_4 = 0.9825211$, and $\mathbf{H}_8(\mathbf{B})$ with $\gamma = \gamma_8 = 1.151922$, respectively, so that the complexity of computing the matrix exponential by this method is bounded by

$$\eta(m^*) = \tfrac{5}{3} + \lceil \kappa(\varepsilon) + \tfrac{1}{2}\log_2(\|\mathbf{A}^2\|) \rceil, \tag{39}$$

where $\kappa(\varepsilon) = m^* - \frac{1}{2}\log_2(\gamma_{2m^*}) = 2.0127$ for $\varepsilon = 2^{-23}$, and $\kappa(\varepsilon) = 3.8980$ for $\varepsilon = 2^{-52}$. In Appendix A2, we compare this algorithm with that based on the better-known direct Padé approximations of the matrix exponential.

3. COMMUTATOR POWER SERIES

The kth *commutator power* $\{\mathbf{V}, \mathbf{A}^k\}$ is defined recursively as

$$\{\mathbf{V}, \mathbf{A}^0\} = \mathbf{V}, \qquad \{\mathbf{V}, \mathbf{A}^k\} = \left[\{\mathbf{V}, \mathbf{A}^{k-1}\}, \mathbf{A}\right] \qquad (k > 0), \tag{40}$$

where $[\mathbf{X}, \mathbf{Y}]$ denotes the commutator of the matrices $\mathbf{X}$ and $\mathbf{Y}$ as usual. Although the notation of Eq. (40) is fairly standard [41, 42, 70], the name is not. Since the commutator powers satisfy the addition property for powers, i.e., $\{\mathbf{V}, \mathbf{A}^{k+l}\} = \{\{\mathbf{V}, \mathbf{A}^k\}, \mathbf{A}^l\}$, this name nevertheless fits the notation and is compatible with the action of the matrix $\mathbf{A}$.

Let $\mathbf{F}(\mathbf{A})$ be an analytic function of a matrix argument with power series representation $\mathbf{F}(\mathbf{A}) = \sum_{k=0}^{\infty} c_k \mathbf{A}^k$. Then for any matrix $\mathbf{V}$ of the same size, the corresponding *commutator series* (series in commutator powers or

commutator function) [42] is given by

$$\{\mathbf{V}, \mathbf{F}(\mathbf{A})\} = \sum_{k=0}^{\infty} c_k \{\mathbf{V}, \mathbf{A}^k\}. \tag{41}$$

The identity $\{\mathbf{V}, \mathbf{F}(-\mathbf{A})\} = \{\mathbf{F}(\mathbf{A}), \mathbf{V}\} = \sum_{k=0}^{\infty} c_k \{\mathbf{A}^k, \mathbf{V}\}$ follows at once.

As we shall see, the first directional derivative of the matrix exponential admits an expansion as in Eq. (41), which immediately raises the question of when such commutator series converge.

3.1. *Convergence Criteria for Commutator Series*

The study of the convergence of commutator series is greatly facilitated by the Kronecker product $\otimes$, sum $\oplus$, and col operations (see, e.g., Refs. [27, 33, 50]) together with the $n^2 \times n^2$ matrix (cf. Ref. [53])

$$\mathbf{K}(\mathbf{A}) = \mathbf{A}^{\mathrm{T}} \oplus (-\mathbf{A}) = \mathbf{A}^{\mathrm{T}} \otimes \mathbf{I} - \mathbf{I} \otimes \mathbf{A}. \tag{42}$$

PROPOSITION 3.1. *The linear operator* $\mathbf{V} \to \{\mathbf{V}, \mathbf{A}^k\}$ *has the matrix representation*

$$\operatorname{col}(\{\mathbf{V}, \mathbf{A}^k\}) = \mathbf{K}^k(\mathbf{A}) \operatorname{col}(\mathbf{V}) \qquad (k \geq 0). \tag{43}$$

Proof. From the well-known fact that $\operatorname{col}(\mathbf{AVB}) = (\mathbf{B}^{\mathrm{T}} \otimes \mathbf{A}) \operatorname{col}(\mathbf{V})$ (see, e.g., Refs. [6, 27]), we obtain $\operatorname{col}(\{\mathbf{V}, \mathbf{A}\}) = \mathbf{K}(\mathbf{A}) \operatorname{col}(\mathbf{V})$. The rest follows from the recursion (40), since $\operatorname{col}(\{\mathbf{V}, \mathbf{A}^k\}) = \mathbf{K}(\mathbf{A}) \operatorname{col}(\{\mathbf{V}, \mathbf{A}^{k-1}\}) = \cdots = \mathbf{K}^k(\mathbf{A}) \operatorname{col}(\mathbf{V})$. ∎

As is well known, the mn eigenvalues of the Kronecker sum of two square matrices $\mathbf{B}$ and $\mathbf{C}$ are given in terms of their eigenvalues by $\mu_{ij} = \lambda_i(\mathbf{B}) + \lambda_j(\mathbf{C})$ $(i = 1, \ldots, m$ and $j = 1, \ldots, n)$. Thus the spectral radius of $\mathbf{K}(\mathbf{A})$ is given by

$$\rho(\mathbf{K}(\mathbf{A})) = \max_{i,j} |\lambda_i(\mathbf{A}) - \lambda_j(\mathbf{A})| = \sigma(\mathbf{A}). \tag{44}$$

The quantity $\sigma(\mathbf{A})$ is known as the *spread* of $\mathbf{A}$ [43]. The following result shows that the spread plays the same role in the convergence of commutator series as the spectral radius in the convergence of power series.

THEOREM 3.2. *Let* $\mathbf{F}(\mathbf{A})$ *be an analytic function of a matrix argument which converges for all matrices satisfying* $\rho(\mathbf{A}) < \xi$. *Then the commutator series* $\{\mathbf{V}, \mathbf{F}(\mathbf{A})\}$ *converges provided that* $\sigma(\mathbf{A}) < \xi$, *and is given by*

$$\operatorname{col}(\{\mathbf{V}, \mathbf{F}(\mathbf{A})\}) = \mathbf{F}(\mathbf{K}(\mathbf{A})) \operatorname{col}(\mathbf{V}). \tag{45}$$

Proof. By Eq. (43), acting on the mth partial sum in Eq. (41) with the col operator results in a matrix polynomial in $\mathbf{K}(\mathbf{A})$, namely

$(\sum_{k=0}^{m} c_k \mathbf{K}^k(\mathbf{A}))\mathrm{col}(\mathbf{V})$. This converges as $m \to \infty$ whenever $\rho(\mathbf{K}(\mathbf{A})) < \xi$ by a well-known result on matrix power series, which in view of (44) implies the convergence of $\{\mathbf{V}, \mathbf{F}(\mathbf{A})\}$ whenever $\sigma(\mathbf{A}) < \xi$ and hence Eq. (45) as well. ∎

The simplest bound on $\sigma(\mathbf{A})$ follows from the definition in Eq. (44), namely $\sigma(\mathbf{A}) \leq 2\rho(\mathbf{A})$, which implies that the radius of convergence of $\mathbf{F}(\mathbf{K}(\mathbf{A}))$ may be only half that of $\mathbf{F}(\mathbf{A})$. If, however, $\mathbf{A}$ is a strictly (un)stable matrix, then $\sigma(\mathbf{A}) \leq \rho(\mathbf{A})$. Other, sharper bounds on $\sigma(\mathbf{A})$ may be found in [43].

We mention that if one defines the *anticommutator* of two square matrices of the same size as $[\mathbf{X}, \mathbf{Y}]_+ = \mathbf{XY} + \mathbf{YX}$, then most of the above definitions and results can be extended in a straightforward way to anti-commutator powers and series.

3.2. *The Evaluation of Commutator Series*

In some cases it is possible to evaluate a commutator series exactly by converting it to an ordinary power series in $\mathbf{K}(\mathbf{A})$, although the computational complexity of this procedure is generally at least $O(n^6)$. For example, the geometric commutator series $\{\mathbf{V}, \mathbf{G}(\mathbf{A})\} = \sum_{k=0}^{\infty} \{\mathbf{V}, \mathbf{A}^k\}$ can be summed as

$$\mathrm{col}(\{\mathbf{V}, \mathbf{G}(\mathbf{A})\}) = \sum_{k=0}^{\infty} \mathbf{K}^k(A)\mathrm{col}(\mathbf{V}) = (\mathbf{I} - \mathbf{K}(\mathbf{A}))^{-1}\mathrm{col}(\mathbf{V}) \quad (46)$$

provided that $\sigma(\mathbf{A}) < 1$. Another example is the exponential commutator series, i.e.,

$$\{\mathbf{V}, e^{\mathbf{A}}\} = \sum_{k=0}^{\infty} \frac{\{\mathbf{V}, \mathbf{A}^k\}}{k!}. \quad (47)$$

Applying the col operation to both sides of this equation and using the Kronecker addition formula $\exp(\mathbf{A}^{\mathrm{T}} \oplus (-\mathbf{A})) = \exp(\mathbf{A}^{\mathrm{T}}) \otimes \exp(-\mathbf{A})$ [27] result in

$$\begin{aligned} \mathrm{col}(\{\mathbf{V}, e^{\mathbf{A}}\}) &= e^{\mathbf{K}(\mathbf{A})}\mathrm{col}(\mathbf{V}) = e^{\mathbf{A}^{\mathrm{T}} \oplus (-\mathbf{A})}\mathrm{col}(\mathbf{V}) \\ &= (e^{\mathbf{A}^{\mathrm{T}}} \otimes e^{-\mathbf{A}})\mathrm{col}(\mathbf{V}) = \mathrm{col}(e^{-\mathbf{A}}\mathbf{V}e^{\mathbf{A}}). \end{aligned} \quad (48)$$

This proves the well-known Lie group relation between the commutator exponential and the adjoint operator that was first derived by Baker and Hausdorff [5, 29], namely $\{\mathbf{V}, \exp(\mathbf{A})\} = \exp(-\mathbf{A})\mathbf{V}\exp(\mathbf{A})$. It also shows that this particular commutator series can be evaluated in only $O(n^3)$ time.

In most cases, however, it is more efficient to approximate the commutator series (41) directly, by truncating the series so that the bound on the norm of remainder is less than the desired precision. This will be illustrated in Section 5.1. A commutator polynomial of degree ν,

$$\Gamma_{\mathbf{V}}^{\nu}(\mathbf{A}) = \sum_{k=0}^{\nu} a_k\{\mathbf{V}, \mathbf{A}^k\}, \tag{49}$$

can be evaluated efficiently by the following algorithm, which constitutes a commutator version of Horner's scheme:

$$\begin{aligned} &\textit{Set}\colon\ \Gamma_{\nu} = a_{\nu}\mathbf{V} \\ &\textit{Loop}\colon\ k = \nu - 1, \ldots, 0 \\ &\textit{End}\colon\ \Gamma_k = a_k\mathbf{V} + [\Gamma_{k+1}, \mathbf{A}] \end{aligned} \tag{50}$$

In the case of an even commutator polynomial $\Gamma_{\mathbf{V}}^{2\nu}(\mathbf{A}) = \sum_{k=0}^{\nu} a_{2k}\{\mathbf{V}, \mathbf{A}^{2k}\}$, this algorithm is modified simply by replacing the single commutator with a double commutator. Two observations will be relevant later on. First, since $\{\mathbf{V}, \mathbf{A}^2\} \neq [\mathbf{V}, \mathbf{A}^2]$ in general, it is not possible to halve the number of matrix multiplications for even commutator polynomials, as is the case for ordinary matrix polynomials. Second, if $\mathbf{A}$ and $\mathbf{V}$ are symmetric matrices, then the odd commutator powers are skew-symmetric while the even ones are symmetric, so the time and storage required can be halved in the symmetric case.

4. THE FIRST DIRECTIONAL DERIVATIVE OF THE MATRIX EXPONENTIAL

Like the matrix exponential itself, its first directional derivative admits a variety of representations, which in turn reveal a number of useful properties. Although some of these have been rediscovered numerous times, a reasonably unified presentation has been lacking. In Section 4.1, we present some elementary properties of the directional derivative, most of which are consequences of the integral representation (2). Then in Section 4.2, we derive several new compact spectral representations of the first directional derivative, and finally in Section 4.3 we give a number of different series and closed form representations in terms of Kronecker products. Whenever it is convenient, these results are extended to arbitrary analytic functions of a matrix argument.

4.1. *Properties of the Directional Derivative*

We begin by summarizing some of the most important properties of the first directional derivative.

PROPOSITION 4.1. *If* $\mathbf{D}_{\mathbf{V}}(t,\mathbf{A})$ *denotes the first directional derivative of the matrix exponential, then*

(a) $\mathbf{D}_{\mathbf{V}}(t,\mathbf{A}) = \mathbf{D}_{t\mathbf{V}}(1,t\mathbf{A})$.

(b) *If* $\mathbf{A}$ *and* $\mathbf{V}$ *are Hermitian matrices, then so is* $\mathbf{D}_{\mathbf{V}}(t,\mathbf{A})$.

(c) *If* $\mathbf{A}$ *and* $\mathbf{V}$ *are skew-Hermitian matrices, then*

$$\mathbf{D}_{\mathbf{V}}^{\mathrm{H}}(1,\mathbf{A}) = -e^{-\mathbf{A}}\mathbf{D}_{\mathbf{V}}(1,\mathbf{A})e^{-\mathbf{A}} = \mathbf{D}_{-\mathbf{V}}(1,-\mathbf{A}) = \mathbf{D}_{\mathbf{V}}(-1,\mathbf{A}). \quad (51)$$

(d) *Under the similarity transformation* $\mathbf{A} = \mathbf{T}\mathbf{B}\mathbf{T}^{-1}$, *the directional derivative is transformed to*

$$\mathbf{D}_{\mathbf{V}}(1,\mathbf{A}) = \mathbf{T}\mathbf{D}_{\mathbf{W}}(1,\mathbf{B})\mathbf{T}^{-1}, \qquad \text{where } \mathbf{W} = \mathbf{T}^{-1}\mathbf{V}\mathbf{T}. \quad (52)$$

(e) *For any two square matrices* $\mathbf{A}$ *and* $\mathbf{V}$, *the following addition formula holds*:

$$\mathbf{D}_{\mathbf{V}}(t+s,\mathbf{A}) = e^{s\mathbf{A}}\mathbf{D}_{\mathbf{V}}(t,\mathbf{A}) + \mathbf{D}_{\mathbf{V}}(s,\mathbf{A})e^{t\mathbf{A}}. \quad (53)$$

(f) *For* $t = s$, *the addition formula becomes the doubling formula*

$$\mathbf{D}_{\mathbf{V}}(2t,\mathbf{A}) = e^{t\mathbf{A}}\mathbf{D}_{\mathbf{V}}(t,\mathbf{A}) + \mathbf{D}_{\mathbf{V}}(t,\mathbf{A})e^{t\mathbf{A}}. \quad (54)$$

(g) *Repeated use of the doubling formula leads to a "two-sided" analogue of the iterative squaring formula for the matrix exponential, namely*

$$\mathbf{D}_{\mathbf{V}}(2^{m+1}t,\mathbf{A}) = e^{2^m t\mathbf{A}}\mathbf{D}_{\mathbf{V}}(2^m t,\mathbf{A}) + \mathbf{D}_{\mathbf{V}}(2^m t,\mathbf{A})e^{2^m t\mathbf{A}}. \quad (55)$$

(h) *For* $t = -s$, *the addition formula yields the following analogue of the inversion formula for the matrix exponential*:

$$\mathbf{D}_{\mathbf{V}}(-t,\mathbf{A}) = -e^{-t\mathbf{A}}\mathbf{D}_{\mathbf{V}}(t,\mathbf{A})e^{-t\mathbf{A}}. \quad (56)$$

Proof. Statements (a) through (e) follow from the integral representation (2) by straightforward calculus, while (f) through (h) are immediate consequences of the addition formula (e). ∎

Most of the above properties of $\mathbf{D}_{\mathbf{V}}(t,\mathbf{A})$ also hold for its discrete analogue, the power directional derivative $\mathbf{D}_{\mathbf{V}}(\mathbf{A}^p)$, although the proofs must be adapted to use summation and the discrete variable p, rather than

integration and the continuous variable t. For example, it is easy to show (by induction) that the addition formula

$$\mathbf{D}_{\mathbf{V}}(\mathbf{A}^{p+q}) = \mathbf{A}^{p}\mathbf{D}_{\mathbf{V}}(\mathbf{A}^{q}) + \mathbf{D}_{\mathbf{V}}(\mathbf{A}^{p})\mathbf{A}^{q} \tag{57}$$

holds. For $p = q$ this yields a doubling formula

$$\mathbf{D}_{\mathbf{V}}(\mathbf{A}^{2p}) = \mathbf{A}^{p}\mathbf{D}_{\mathbf{V}}(\mathbf{A}^{p}) + \mathbf{D}_{\mathbf{V}}(\mathbf{A}^{p})\mathbf{A}^{p}. \tag{58}$$

Differentiation of Eq. (2) shows that $\mathbf{D}_{\mathbf{V}}(t,\mathbf{A})$ satisfies both a left and right inhomogeneous system of differential equations, i.e.,

$$\frac{d}{dt}\mathbf{D}_{\mathbf{V}}(t,\mathbf{A}) = \begin{cases} \mathbf{A}\mathbf{D}_{\mathbf{V}}(t,\mathbf{A}) + \mathbf{V}e^{t\mathbf{A}} \\ \mathbf{D}_{\mathbf{V}}(t,\mathbf{A})\mathbf{A} + e^{t\mathbf{A}}\mathbf{V} \end{cases} \qquad (\mathbf{D}_{\mathbf{V}}(0,\mathbf{A}) = \mathbf{0}). \tag{59}$$

Similarly, differencing its discrete analogue in Eq. (5) shows that $\mathbf{D}_{\mathbf{V}}(\mathbf{A}^{p})$ satisfies two inhomogeneous systems of difference equations

$$\mathbf{D}_{\mathbf{V}}(\mathbf{A}^{p+1}) = \begin{cases} \mathbf{A}\mathbf{D}_{\mathbf{V}}(\mathbf{A}^{p}) + \mathbf{V}\mathbf{A}^{p} \\ \mathbf{D}_{\mathbf{V}}(\mathbf{A}^{p})\mathbf{A} + \mathbf{A}^{p}\mathbf{V} \end{cases} \qquad (\mathbf{D}_{\mathbf{V}}(\mathbf{A}^{0}) = \mathbf{0}). \tag{60}$$

PROPOSITION 4.2. *The exponential and power directional derivatives satisfy the linear commutator equations*

$$[\mathbf{D}_{\mathbf{V}}(t,\mathbf{A}),\mathbf{A}] = [\mathbf{V},e^{t\mathbf{A}}] \tag{61}$$

and

$$[\mathbf{D}_{\mathbf{V}}(\mathbf{A}^{p}),\mathbf{A}] = [\mathbf{V},\mathbf{A}^{p}]. \tag{62}$$

Proof. These equations follow by subtraction of the right-hand sides of Eqs. (59) and (60), respectively. ∎

Observe that if we apply the col operation to both sides of Eq. (61), we get

$$\mathbf{K}(\mathbf{A})\mathrm{col}(\mathbf{D}_{\mathbf{V}}(t,\mathbf{A})) = \mathbf{K}(\exp(\mathbf{A}))\mathrm{col}(\mathbf{V}). \tag{63}$$

Although the matrix $\mathbf{K}(\mathbf{A}) = \mathbf{A}^{\mathrm{T}} \oplus (-\mathbf{A})$ has at least n zero eigenvalues and hence cannot be inverted directly, we know that a unique solution nevertheless exists (indicating that the right-hand side of Eq. (63) must be in the range of $\mathbf{K}(\mathbf{A})$). Equation (63) was also derived by Deif [18], who used the eigensystem of $\mathbf{A}$ to obtain a spectral representation for $\mathbf{D}_{\mathbf{V}}(t,\mathbf{A})$.

We now show that a similar linear equation holds for any analytic function. In what follows $\mathbf{F}(\mathbf{A}) = \sum_{k=0}^{\infty} c_k \mathbf{A}^{k}$ denotes a matrix analytic

function which converges for all matrices of spectral radius $\rho(\mathbf{A}) < \xi$, while f denotes the scalar version of $\mathbf{F}$.

THEOREM 4.3. *If* $\mathbf{D}_{\mathbf{V}}(\mathbf{F}(\mathbf{A}))$ *is the first directional derivative of* $\mathbf{F}(\mathbf{A})$ *in the direction* $\mathbf{V}$, *then it satisfies the linear commutator equation*

$$[\mathbf{D}_{\mathbf{V}}(\mathbf{F}(\mathbf{A})), \mathbf{A}] = [\mathbf{V}, \mathbf{F}(\mathbf{A})], \tag{64}$$

or equivalently,

$$\mathbf{K}(\mathbf{A})\mathrm{col}(\mathbf{D}_{\mathbf{V}}(t, \mathbf{A})) = \mathbf{K}(\mathbf{F}(\mathbf{A}))\mathrm{col}(\mathbf{V}). \tag{65}$$

Proof. From the power series expansion of $\mathbf{F}(\mathbf{A})$ it follows that

$$\mathbf{D}_{\mathbf{V}}(\mathbf{F}(\mathbf{A})) = \sum_{k=1}^{\infty} c_k \mathbf{D}_{\mathbf{V}}(\mathbf{A}^k). \tag{66}$$

If we now form the commutator of both sides of (66) with $\mathbf{A}$, then by linearity of the commutator together with Eq. (62)

$$\begin{aligned}[\mathbf{D}_{\mathbf{V}}(\mathbf{F}(\mathbf{A})), \mathbf{A}] &= \left[\sum_{k=1}^{\infty} c_k \mathbf{D}_{\mathbf{V}}(\mathbf{A}^k), \mathbf{A}\right] = \sum_{k=1}^{\infty} c_k [\mathbf{D}_{\mathbf{V}}(\mathbf{A}^k), \mathbf{A}] \\ &= \sum_{k=1}^{\infty} c_k [\mathbf{V}, \mathbf{A}^k] = \left[\mathbf{V}, \sum_{k=1}^{\infty} c_k \mathbf{A}^k\right] = [\mathbf{V}, \mathbf{F}(\mathbf{A})]. \end{aligned} \tag{67}$$

This proves Eq. (64). If we rewrite both sides of this equation in terms of the matrix representation of a commutator from Eq. (42), then Eq. (65) follows at once. ∎

We now define the generalized *divided difference matrix* of $\mathbf{F}$ at $\mathbf{A}$ by

$$\mathbf{Q}(\mathbf{F}|\mathbf{A}) = \frac{\mathbf{K}(\mathbf{F}(\mathbf{A}))}{\mathbf{K}(\mathbf{A})} = \frac{\mathbf{F}(\mathbf{A}^{\mathrm{T}}) \oplus (-\mathbf{F}(\mathbf{A}))}{\mathbf{A}^{\mathrm{T}} \oplus (-\mathbf{A})}, \tag{68}$$

prove that the division in Eq. (68) is in fact well defined, and show how $\mathbf{Q}(\mathbf{F}|\mathbf{A})$ maps $\mathbf{V}$ into $\mathbf{D}_{\mathbf{V}}(\mathbf{F}(\mathbf{A}))$.

THEOREM 4.4. *Equation* (65) *has the unique solution*

$$\mathrm{col}(\mathbf{D}_{\mathbf{V}}(\mathbf{F}(\mathbf{A}))) = \mathbf{Q}(\mathbf{F}|\mathbf{A})\mathrm{col}(\mathbf{V}), \tag{69}$$

where $\mathbf{Q}(\mathbf{F}|\mathbf{A})$ *exists for all matrices satisfying* $\rho(\mathbf{A}) < \xi$, *and is given by*

$$\mathbf{Q}(\mathbf{F}|\mathbf{A}) = \frac{\mathbf{K}(\mathbf{F}(\mathbf{A}))}{\mathbf{K}(\mathbf{A})} = \sum_{k=1}^{\infty} c_k \sum_{j=1}^{k} (\mathbf{A}^{\mathrm{T}})^{k-j} \otimes \mathbf{A}^{j-1}. \tag{70}$$

Proof. First, the vanishing of the commutator $[\mathbf{K}(\mathbf{A}), \mathbf{K}(\mathbf{F}(\mathbf{A}))]$ follows easily from the fact that $\mathbf{A}$ and $\mathbf{F}(\mathbf{A})$ commute, together with the mixed product rule for the Kronecker product. Next, the matrix $\mathbf{K}(\mathbf{F}(\mathbf{A}))$ is factored into the product $\mathbf{K}(\mathbf{A})\mathbf{Q}(\mathbf{F}|\mathbf{A})$, as follows:

$$\begin{aligned}
\mathbf{K}(\mathbf{F}(\mathbf{A})) &= \mathbf{F}(\mathbf{A}^{\mathrm{T}}) \otimes \mathbf{I} - \mathbf{I} \otimes \mathbf{F}(\mathbf{A}) = \sum_{k=0}^{\infty} c_k (\mathbf{A}^{\mathrm{T}})^k \otimes \mathbf{I} - \mathbf{I} \otimes \sum_{k=0}^{\infty} c_k \mathbf{A}^k \\
&= \sum_{k=0}^{\infty} c_k \left((\mathbf{A}^{\mathrm{T}} \otimes \mathbf{I})^k - (\mathbf{I} \otimes \mathbf{A})^k \right) \\
&= (\mathbf{A}^{\mathrm{T}} \otimes \mathbf{I} - \mathbf{I} \otimes \mathbf{A}) \left(\sum_{k=1}^{\infty} c_k \sum_{j=1}^{k} (\mathbf{A}^{\mathrm{T}})^{k-j} \otimes \mathbf{A}^{j-1} \right) \\
&= \mathbf{K}(\mathbf{A})\mathbf{Q}(\mathbf{F}|\mathbf{A}).
\end{aligned} \tag{71}$$

To show that we can cancel $\mathbf{K}(\mathbf{A})$ on both sides of Eq. (65) regardless of its rank, we compute $\mathbf{D}_{\mathbf{V}}(\mathbf{F}(\mathbf{A}))$ directly from the power series of $\mathbf{F}(\mathbf{A})$, namely

$$\begin{aligned}
\operatorname{col}(\mathbf{D}_{\mathbf{V}}(\mathbf{F}(\mathbf{A}))) &= \operatorname{col}\left(\sum_{k=1}^{\infty} c_k \mathbf{D}_{\mathbf{V}}(\mathbf{A}^k) \right) \\
&= \sum_{k=1}^{\infty} c_k \operatorname{col}\left(\sum_{j=1}^{k} \mathbf{A}^{k-j} \mathbf{V} \mathbf{A}^{j-1} \right) \\
&= \sum_{k=1}^{\infty} c_k \operatorname{col}\left(\sum_{j=1}^{k} \mathbf{A}^{j-1} \mathbf{V} \mathbf{A}^{k-j} \right) \\
&= \sum_{k=1}^{\infty} c_k \left(\sum_{j=1}^{k} (\mathbf{A}^{\mathrm{T}})^{k-j} \otimes \mathbf{A}^{j-1} \right) \operatorname{col}(\mathbf{V}) \\
&= \mathbf{Q}(\mathbf{F}|\mathbf{A}) \operatorname{col}(\mathbf{V}).
\end{aligned} \tag{72}$$

Finally, the existence of $\mathbf{Q}(\mathbf{F}|\mathbf{A})$ follows from the fact that its eigenvalues q_{rs} $(r, s = 1, \ldots, n)$, are given by [24] (cf. Theorem 4.8 below)

$$q_{rs} = \sum_{k=1}^{\infty} c_k \sum_{j=1}^{k} \lambda_r^{k-j} \lambda_s^{j-1} = \begin{cases} (f(\lambda_r) - f(\lambda_s))/(\lambda_r - \lambda_s), & \lambda_r \neq \lambda_s \\ f'(\lambda_r), & \lambda_r = \lambda_s, \end{cases} \tag{73}$$

and all of these divided differences and derivatives are defined whenever $\rho(\mathbf{A}) < \xi$. ∎

This theorem unifies several earlier results [24, 38, 53]. (Note that the expressions in Theorem 4 of Ref. [53] are much too complicated, since they require either infinite series, a commutative inverse of $\mathbf{K}(\mathbf{A})$, or a limit process, whereas in Ref. [24] there is a missing matrix transpose in a Kronecker product.) Although Eq. (69) is theoretically simple, the expression for $\mathbf{Q}(\mathbf{F}|\mathbf{A})$ in Eq. (70) is much too expensive for computational work in general. We shall therefore look for alternative representations.

4.2. *The Spectral Representations*

The use of spectral methods to study the perturbation of solutions to linear ODEs appears to have first been described by Magnus [42]. Spectral representations of $\mathbf{D}_{\mathbf{V}}(t, \mathbf{A})$ have subsequently been presented by several authors working independently in different fields [3, 18, 24, 34, 38, 72]. Here we present a new formula for the first directional derivative in terms of the eigenstructure of a semi-simple matrix $\mathbf{A}$ and the Hadamard (or entry-by-entry) matrix product, which is more concise than the previous formulae.

THEOREM 4.5. *If* $\mathbf{A} = \mathbf{U}\Lambda\mathbf{U}^{-1}$ *is the spectral decomposition of a semi-simple matrix* $\mathbf{A}$, *its directional derivative in the direction* $\mathbf{V}$ *is given by*

$$\mathbf{D}_{\mathbf{V}}(t, \mathbf{A}) = \mathbf{U}\big(\bar{\mathbf{V}} \odot \Phi(t)\big)\mathbf{U}^{-1}, \tag{74}$$

where $\bar{\mathbf{V}} = \mathbf{U}^{-1}\mathbf{V}\mathbf{U}$ *and* $\bar{\mathbf{V}} \odot \Phi(t)$ *denotes the Hadamard product of* $\bar{\mathbf{V}}$ *with the matrix* $\Phi(t)$ *whose entries are given by*

$$\phi_{ij}(t) = \phi_{ji}(t) = \begin{cases} \big(e^{t\lambda_i} - e^{t\lambda_j}\big)/\big(\lambda_i - \lambda_j\big) & \text{if } \lambda_i \neq \lambda_j \\ t\, e^{t\lambda_i} & \text{if } \lambda_i = \lambda_j. \end{cases} \tag{75}$$

Proof. Substituting the spectral decomposition of $\exp(t\mathbf{A})$ into the integral representation (2), we obtain

$$\begin{aligned} \mathbf{D}_{\mathbf{V}}(t, \mathbf{A}) &= \int_0^t \mathbf{U} e^{(t-\tau)\Lambda} \mathbf{U}^{-1} \mathbf{V} \mathbf{U}\, e^{\tau\Lambda} \mathbf{U}^{-1}\, d\tau \\ &= \mathbf{U}\left(\int_0^t e^{(t-\tau)\Lambda} \bar{\mathbf{V}}\, e^{\tau\Lambda}\, d\tau \right) \mathbf{U}^{-1}. \end{aligned} \tag{76}$$

The left and right multiplication of $\bar{\mathbf{V}}$ by diagonal matrices can be expressed in terms of the Hadamard product as

$$e^{(t-\tau)\Lambda} \bar{\mathbf{V}}\, e^{\tau\Lambda} = \bar{\mathbf{V}} \odot \Psi(t, \tau), \tag{77}$$

where $\Psi(t,\tau)$ is a matrix whose entries are $\psi_{ij}(t,\tau) = \exp((t-\tau)\lambda_j + \tau\lambda_i)$. Since $\bar{\mathbf{V}}$ is a constant matrix, $\int_0^t \bar{\mathbf{V}}\odot\Psi(t,\tau)\,d\tau = \bar{\mathbf{V}}\odot\int_0^t \Psi(t,\tau)\,d\tau = \bar{\mathbf{V}}\odot\Phi(t)$. ∎

We note that each entry $\phi_{ij}(t)$ also satisfies addition, doubling, and inversion formulae similar to $\mathbf{D}_{\mathbf{V}}(t,\mathbf{A})$.

By placing the origin at the midpoint of the integration interval in Eq. (2), we obtain the time-symmetric integral representation of $\mathbf{D}_{\mathbf{V}}(t,\mathbf{A})$

$$\mathbf{D}_{\mathbf{V}}(t,\mathbf{A}) = te^{t\mathbf{A}/2}\left(\frac{1}{t}\int_{-t/2}^{t/2} e^{-\tau\mathbf{A}}\mathbf{V}e^{\tau\mathbf{A}}\,d\tau\right)e^{t\mathbf{A}/2}. \tag{78}$$

This expresses the directional derivative in terms of the *integral average* of $\mathbf{V}$, and leads to a different spectral representation.

THEOREM 4.6. *If* $\mathbf{A} = \mathbf{U}\Lambda\mathbf{U}^{-1}$ *is the spectral decomposition of a semi-simple matrix* $\mathbf{A}$, *its directional derivative in the direction* $\mathbf{V}$ *is given by*

$$\begin{aligned}\mathbf{D}_{\mathbf{V}}(t,\mathbf{A}) &= te^{t\mathbf{A}/2}\mathbf{U}\big(\bar{\mathbf{V}}\odot\Theta(t)\big)\mathbf{U}^{-1}e^{t\mathbf{A}/2}\\ &= t\mathbf{U}\big(e^{t\Lambda/2}\big(\bar{\mathbf{V}}\odot\Theta(t)\big)e^{t\Lambda/2}\big)\mathbf{U}^{-1},\end{aligned} \tag{79}$$

where $\bar{\mathbf{V}} = \mathbf{U}^{-1}\mathbf{V}\mathbf{U}$ *and the matrix* $\mathbf{Q}(t)$ *has the entries*

$$\theta_{ij}(t) = \theta_{ji}(t) = \begin{cases}\dfrac{\sinh\big(t(\lambda_i-\lambda_j)/2\big)}{t(\lambda_i-\lambda_j)/2} & \text{if } \lambda_i \neq \lambda_j\\ 1 & \text{if } \lambda_i = \lambda_j.\end{cases} \tag{80}$$

Proof. On substituting the spectral decomposition of $\exp(t\mathbf{A})$ into the integral of Eq. (78) and simplying much as in Eq. (76), we obtain

$$\mathbf{D}_{\mathbf{V}}(t,\mathbf{A}) = t\mathbf{U}e^{t\Lambda/2}\left(\frac{1}{t}\int_{-t/2}^{t/2} e^{-\tau\Lambda}\bar{\mathbf{V}}e^{\tau\Lambda}\,d\tau\right)e^{t\Lambda/2}\mathbf{U}^{-1}. \tag{81}$$

As in the proof of Theorem 4.5, the left and right multiplication of $\bar{\mathbf{V}}$ by diagonal matrices can be expressed as $e^{-\tau\Lambda}\bar{\mathbf{V}}e^{\tau\Lambda} = \bar{\mathbf{V}}\odot\Xi(\tau)$, where $\Xi(\tau)$ is the matrix whose entries are given by $\xi_{ij}(\tau) = \exp(\tau(-\lambda_i+\lambda_j))$. The elementary integration of $\xi_{ij}(\tau)$ between the indicated limits generates the functions $\theta_{ij}(t)$ in Eq. (80). ∎

The scalar functions in Eq. (80) are numerically more stable than the functions in Eq. (75). In fact, numerical stabilization of the difference quotients $\phi_{ij} = (e^{t\lambda_i} - e^{t\lambda_j})/(\lambda_i - \lambda_j)$ for small values of $|\lambda_i - \lambda_j|$ leads to

precisely the functions θ_{ij} in Eq. (80), while $\sinh(z)/z$ may be evaluated by an even polynomial or a rational approximation for small values of $|z|$.

The spectral representation of $\mathbf{D}_{\mathbf{V}}(\mathbf{A}^p)$ is given by the next theorem.

THEOREM 4.7. *If* $\mathbf{A} = \mathbf{U}\Lambda\mathbf{U}^{-1}$ *is the spectral representation of a semisimple matrix* $\mathbf{A}, \mathbf{V}$ *is an arbitrary square matrix of the same size, and* $\bar{\mathbf{V}} = \mathbf{U}^{-1}\mathbf{V}\mathbf{U}$, *then*

$$\mathbf{D}_{\mathbf{V}}(\mathbf{A}^p) = \mathbf{U}\big(\bar{\mathbf{V}} \odot \tilde{\Phi}(p)\big)\mathbf{U}^{-1}, \tag{82}$$

where $\tilde{\Phi}(p) = \{\tilde{\phi}_{ij}(p)\}$ *is defined as*

$$\tilde{\phi}_{ij}(p) = \tilde{\phi}_{ji}(p) = \begin{cases} (\lambda_i^p - \lambda_j^p)/(\lambda_i - \lambda_j) & \text{if } \lambda_i \neq \lambda_j \\ p\lambda_i^{p-1} & \text{if } \lambda_i = \lambda_j. \end{cases} \tag{83}$$

Proof. Substitution of $\mathbf{A}^k = \mathbf{U}\Lambda^k\mathbf{U}^{-1}$ into Eq. (5) results in

$$\begin{aligned} \mathbf{D}_{\mathbf{V}}(\mathbf{A}^p) &= \sum_{k=1}^{p} \mathbf{A}^{p-k}\mathbf{V}\mathbf{A}^{k-1} = \sum_{k=1}^{p} \mathbf{U}\Lambda^{p-k}\mathbf{U}^{-1}\mathbf{V}\mathbf{U}\Lambda^{k-1}\mathbf{U}^{-1} \\ &= \mathbf{U}\left(\sum_{k=1}^{p} \Lambda^{p-k}\bar{\mathbf{V}}\Lambda^{k-1}\right)\mathbf{U}^{-1}. \end{aligned} \tag{84}$$

The left and right multiplication of $\bar{\mathbf{V}}$ by diagonal matrices can be expressed as $\Lambda^{p-k}\bar{\mathbf{V}}\Lambda^{k-1} = \bar{\mathbf{V}} \odot \tilde{\Psi}(p,k)$, where $\tilde{\psi}_{ij}(p,k) = \lambda_j^{p-k}\lambda_i^{(k-1)}$ are the entries of the $\tilde{\Psi}(p,k)$. Since $\bar{\mathbf{V}}$ is a constant matrix, $\sum_{k=1}^{p} \bar{\mathbf{V}} \odot \tilde{\Psi}(p,k) = \bar{\mathbf{V}} \odot \sum_{k=1}^{p} \tilde{\Psi}(p,k) = \bar{\mathbf{V}} \odot \tilde{\Phi}(p)$, where the last equality follows from the identity $z^p - w^p = (z-w)\sum_{k=1}^{p} z^{p-k}w^{k-1}$ applied to each entry of $\tilde{\Psi}(p,k)$. ∎

Finally, we present a spectral representation for the first directional derivative $\mathbf{D}_{\mathbf{V}}(\mathbf{F}(\mathbf{A}))$ of an analytic matrix function $\mathbf{F}(\mathbf{A}) = \sum_{k=0}^{\infty} c_k\mathbf{A}^k$ in the direction $\mathbf{V}$.

THEOREM 4.8. *If* $\mathbf{A} = \mathbf{U}\Lambda\mathbf{U}^{-1}$ *is the spectral representation of a semisimple matrix* $\mathbf{A}$, $\mathbf{V}$ *is an arbitrary square matrix of the same size, and* $\bar{\mathbf{V}} = \mathbf{U}^{-1}\mathbf{V}\mathbf{U}$, *then*

$$\mathbf{D}_{\mathbf{V}}(\mathbf{F}(\mathbf{A})) = \mathbf{U}\big(\bar{\mathbf{V}} \odot \Delta_{\lambda}(\mathbf{F})\big)\mathbf{U}^{-1}, \tag{85}$$

where

$$\Delta_{\lambda}(\mathbf{F}) = \begin{cases} \big(f(\lambda_i) - f(\lambda_j)\big)/(\lambda_i - \lambda_j) & \text{if } \lambda_i \neq \lambda_j \\ f'(\lambda_i) & \text{if } \lambda_i = \lambda_j \end{cases} \tag{86}$$

is a symmetric divided difference matrix, and f is the scalar version of the function **F**.

Proof. By the linearity of the directional derivative together with Theorem 4.7, we can write

$$\mathbf{D}_{\mathbf{V}}(\mathbf{F}(\mathbf{A})) = \sum_{k=1}^{\infty} c_k \mathbf{D}_{\mathbf{V}}(\mathbf{A}^k) = \sum_{k=1}^{\infty} c_k \mathbf{U}\big(\bar{\mathbf{V}} \odot \tilde{\Phi}(k)\big)\mathbf{U}^{-1}$$

$$= \mathbf{U}\left(\bar{\mathbf{V}} \odot \left(\sum_{k=1}^{\infty} c_k \tilde{\Phi}(k)\right)\right)\mathbf{U}^{-1} = \mathbf{U}\big(\bar{\mathbf{V}} \odot \Delta_{\lambda}(\mathbf{F})\big)\mathbf{U}^{-1}. \quad \blacksquare \quad (87)$$

The matrix $\Delta_{\lambda}(\mathbf{F})$ is sometimes called the Loewner matrix. As pointed out in Ref. [24], it contains the eigenvalues of the first directional derivative of any analytic matrix function $\mathbf{F}(\mathbf{A})$, considered as an operator acting on the direction $\mathbf{V}$. This theorem bears an interesting resemblance to the Daleckiĭ–Krein formula for the derivative of a Hermitian matrix-valued differentiable function $\mathbf{F}(t)$ with respect to scalar t [32].

4.3. *Series Representations and Their Closed Form Solutions*

The Taylor series expansion for $\mathbf{D}_{\mathbf{V}}(t, \mathbf{A})$ relates it to the power directional derivatives $\mathbf{D}_{\mathbf{V}}(\mathbf{A}^k)$ [24, 42, 53]. Here, we offer a concise alternative proof of this result:

THEOREM 4.9. *The Taylor series expansion for* $\mathbf{D}_{\mathbf{V}}(t, \mathbf{A})$ *at* $t = 0$ *is given by*

$$\mathbf{D}_{\mathbf{V}}(t, \mathbf{A}) = \sum_{k=0}^{\infty} \frac{t^k}{k!} \frac{d^k}{dt^k} \mathbf{D}_{\mathbf{V}}(t, \mathbf{A})\bigg|_{t=0}$$

$$= \sum_{k=1}^{\infty} \frac{t^k}{k!} \mathbf{D}_{\mathbf{V}}(\mathbf{A}^k). \quad (88)$$

Proof. Let $\mathbf{Z}_k(t) = (d^k/dt^k)\mathbf{D}_{\mathbf{V}}(t, \mathbf{A})$. By differentiating the integral representation (2) k times, we obtain a difference equation for these derivatives at any time t, i.e.,

$$\mathbf{Z}_{k+1}(t) = \mathbf{A}\mathbf{Z}_k(t) + \mathbf{V}e^{t\mathbf{A}}\mathbf{A}^k. \quad (89)$$

Since $\mathbf{Z}_0(0) = \mathbf{0}$ and the difference equation (89) becomes equivalent to Eq. (60) at $t = 0$, it follows from the uniqueness of the solutions to linear difference equations that the derivatives $\mathbf{Z}_k(0)$ are identical to the power directional derivatives. ∎

The following results are motivated by several earlier representations of a function of a triangular or block triangular matrix and their relations to its derivatives, divided differences, and integrals [16, 35, 47, 57, 62]. For example, Theorem 4.9 enables us to derive the following *functional representation* of the directional derivative:

PROPOSITION 4.10. *If* $\mathbf{B}$ *is the* $2n \times 2n$ *block matrix* $\mathbf{B} = \begin{bmatrix} \mathbf{A} & \mathbf{V} \\ \mathbf{0} & \mathbf{A} \end{bmatrix}$, *then*

$$e^{t\mathbf{B}} = \begin{bmatrix} \exp(t\mathbf{A}) & \mathbf{D}_{\mathbf{V}}(t,\mathbf{A}) \\ \mathbf{0} & \exp(t\mathbf{A}) \end{bmatrix}. \tag{90}$$

Proof. By induction on $\mathbf{B}^k = \mathbf{B}\mathbf{B}^{k-1}$ and the addition formula (57), it follows that the upper off-diagonal block of the matrix $\mathbf{B}^k$ is the kth power directional derivative

$$\mathbf{B}^k = \begin{bmatrix} \mathbf{A}^k & \mathbf{D}_{\mathbf{V}}(\mathbf{A}^k) \\ \mathbf{0} & \mathbf{A}^k \end{bmatrix}. \tag{91}$$

Thus the Taylor series expansion of $\exp(t\mathbf{B})$ may be written as

$$\begin{aligned} e^{t\mathbf{B}} &= \sum_{k=0}^{\infty} \frac{t^k}{k!} \begin{bmatrix} \mathbf{A} & \mathbf{V} \\ \mathbf{0} & \mathbf{A} \end{bmatrix}^k \\ &= \begin{bmatrix} \sum_{k=0}^{\infty} (t\mathbf{A})^k/k! & \sum_{k=1}^{\infty} t^k \mathbf{D}_{\mathbf{V}}(\mathbf{A}^k)/k! \\ \mathbf{0} & \sum_{k=0}^{\infty} (t\mathbf{A})^k/k! \end{bmatrix}. \end{aligned} \tag{92}$$

The proof then follows at once from the Taylor series for $\mathbf{D}_{\mathbf{V}}(t,\mathbf{A})$ in Theorem 4.9. ∎

The explicit appearance of the power directional derivatives in Eqs. (91) and (92) enables us to generalize this result to any analytic function.

THEOREM 4.11. *If* $\mathbf{F}$ *is an analytic matrix function and* $\mathbf{B} = \begin{bmatrix} \mathbf{A} & \mathbf{V} \\ \mathbf{0} & \mathbf{A} \end{bmatrix}$, *then*

$$\mathbf{F}(\mathbf{B}) = \begin{bmatrix} \mathbf{F}(\mathbf{A}) & \mathbf{D}_{\mathbf{V}}(\mathbf{F}(\mathbf{A})) \\ \mathbf{0} & \mathbf{F}(\mathbf{A}) \end{bmatrix}. \tag{93}$$

The proof is almost identical to the proof of Proposition 4.10 and is therefore omitted.

The unusually simple generalization to higher-order directional derivatives is obtained by forming the bidiagonal block matrix

$$\mathbf{B} = \begin{bmatrix} \mathbf{A} & \mathbf{V} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{A} & \mathbf{V} & \cdots & \mathbf{0} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{0} & \cdots & \cdots & \mathbf{A} & \mathbf{V} \\ \mathbf{0} & \mathbf{0} & \cdots & \cdots & \mathbf{A} \end{bmatrix}, \tag{94}$$

where the number of $\mathbf{V}$ blocks along the first superdiagonal is equal to the highest-order derivative q desired. We shall call $\mathbf{B}$ the qth order *directional Jordan block* defined by $\mathbf{A}$ and $\mathbf{V}$.

PROPOSITION 4.12. *The pth power of the qth order directional Jordan block* $\mathbf{B}^p$ *contains the first q directional derivatives of* $\mathbf{A}^p$ *in the direction of the matrix* $\mathbf{V}$, *arranged as follows*:

$$\mathbf{B}^p = \begin{bmatrix} \mathbf{A}^p & \mathbf{D}_{\mathbf{V}}^{[1]}(\mathbf{A}^p)/1! & \cdots & \mathbf{D}_{\mathbf{V}}^{[q]}(\mathbf{A}^p)/q! \\ \mathbf{0} & \mathbf{A}^p & \cdots & \mathbf{D}_{\mathbf{V}}^{[q-1]}(\mathbf{A}^p)/(q-1)! \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{A}^p \end{bmatrix}. \tag{95}$$

Proof. Since $\mathbf{B}$ is block semi-circulant, its powers are block semi-circulant too [17]. This means that $\mathbf{B}^p$ is fully determined by its first block row. Each matrix coefficient in the noncommutative expansion

$$(\mathbf{A} + h\mathbf{V})^p = \mathbf{C}_0^{[p]} + h\mathbf{C}_1^{[p]} + h^2\mathbf{C}_2^{[p]} + \cdots + h^p\mathbf{C}_p^{[p]} \tag{96}$$

is a sum of $\binom{p}{k}$ interlaced products of $\mathbf{A}$ and $\mathbf{V}$ of degree k in $\mathbf{V}$ and $p-k$ in $\mathbf{A}$. For example, if $p = 3$ and $k = 2$ then $\mathbf{C}_2^{[3]} = \mathbf{A}\mathbf{V}^2 + \mathbf{V}\mathbf{A}\mathbf{V} + \mathbf{V}^2\mathbf{A}$. These coefficients are related to the directional derivatives as

$$\mathbf{C}_k^{[p]} = \frac{d^k/dh^k(\mathbf{A} + h\mathbf{V})^p}{k!}\bigg|_{h=0} = \frac{\mathbf{D}_{\mathbf{V}}^{[k]}(\mathbf{A}^p)}{k!}, \tag{97}$$

and in particular $\mathbf{C}_0^{[p]} = \mathbf{A}^p$. The directional derivatives, in turn, are linear operators which satisfy $\mathbf{D}_{\mathbf{V}}^{[k]}(\mathbf{A}^p) = \mathbf{D}_{\mathbf{V}}^{[k-1]}(\mathbf{D}_{\mathbf{V}}^{[1]}(\mathbf{A}^p))$. The substitution of Eq. (57) into this last expression results in

$$\begin{aligned} \mathbf{D}_{\mathbf{V}}^{[k]}(\mathbf{A}^p) &= \mathbf{D}_{\mathbf{V}}^{[k-1]}\big(\mathbf{V}\mathbf{A}^{p-1} + \mathbf{A}\mathbf{D}_{\mathbf{V}}^{[1]}(\mathbf{A}^{p-1})\big) \\ &= \mathbf{V}\mathbf{D}_{\mathbf{V}}^{[k-1]}(\mathbf{A}^{p-1}) + \mathbf{D}_{\mathbf{V}}^{[k-1]}\big(\mathbf{A}\mathbf{D}_{\mathbf{V}}^{[1]}(\mathbf{A}^{p-1})\big). \end{aligned} \tag{98}$$

By repeatedly splitting the operator $\mathbf{D}_{\mathbf{V}}^{[k-1]} = \mathbf{D}_{\mathbf{V}}^{[k-2]}\mathbf{D}_{\mathbf{V}}^{[1]}$ and using the rule for differentiating the product of two matrix functions (in the same direction $\mathbf{V}$), namely

$$\mathbf{D}_{\mathbf{V}}^{[1]}(\mathbf{F}(\mathbf{A})\mathbf{G}(\mathbf{A})) = \mathbf{D}_{\mathbf{V}}^{[1]}(\mathbf{F}(\mathbf{A}))\mathbf{G}(\mathbf{A}) + \mathbf{F}(\mathbf{A})\mathbf{D}_{\mathbf{V}}^{[1]}(\mathbf{G}(\mathbf{A})), \tag{99}$$

we obtain

$$\begin{aligned}\mathbf{D}_{\mathbf{V}}^{[k]}(\mathbf{A}^p) &= \mathbf{V}\mathbf{D}_{\mathbf{V}}^{[k-1]}(\mathbf{A}^{p-1}) + \mathbf{D}_{\mathbf{V}}^{[k-1]}\big(\mathbf{A}\mathbf{D}_{\mathbf{V}}^{[1]}(\mathbf{A}^{p-1})\big)\\ &= 2\mathbf{V}\mathbf{D}_{\mathbf{V}}^{[k-1]}(\mathbf{A}^{p-1}) + \mathbf{D}_{\mathbf{V}}^{[k-2]}\big(\mathbf{A}\mathbf{D}_{\mathbf{V}}^{[2]}(\mathbf{A}^{p-1})\big)\\ &= \cdots\\ &= k\mathbf{V}\mathbf{D}_{\mathbf{V}}^{[k-1]}(\mathbf{A}^{p-1}) + \mathbf{A}\mathbf{D}_{\mathbf{V}}^{[k]}(\mathbf{A}^{p-1}).\end{aligned} \tag{100}$$

By dividing both sides by $k!$ and using the relation in Eq. (97), we obtain the following double recurrence relation between the scaled derivatives and the powers:

$$\mathbf{C}_k^{[p]} = \mathbf{V}\mathbf{C}_{k-1}^{[p-1]} + \mathbf{A}\mathbf{C}_k^{[p-1]}. \tag{101}$$

(A similar "left" recurrence relation, $\mathbf{C}_k^{[p]} = \mathbf{C}_{k-1}^{[p-1]}\mathbf{V} + \mathbf{C}_k^{[p-1]}\mathbf{A}$, can be obtained by starting from the "left" version of Eq. (57), $\mathbf{A}^{p-1}\mathbf{V} + \mathbf{D}_{\mathbf{V}}^{[1]}(\mathbf{A}^{p-1})\mathbf{A}$.)

It follows from Eq. (97) that $\mathbf{C}_k^{[p]} = \mathbf{0}$ whenever $k > p$, and that $\mathbf{C}_p^{[p]} = \mathbf{V}^p$. Thus at $p = 1$, the matrix $\mathbf{B}$ in Eq. (94) clearly displays all the coefficients $\mathbf{C}_k^{[1]}(\mathbf{A})$ $(k = 0, \ldots, q)$ in the first block row. Assuming now that the first block row of $\mathbf{B}^{p-1}$ contains all the coefficients $\mathbf{C}_k^{[p-1]}$ $(k = 0, \ldots, q)$, we find that each off-diagonal block in the first row of $\mathbf{B}\mathbf{B}^{p-1}$ explicitly contains the recurrence of Eq. (101), which completes the induction step. ∎

THEOREM 4.13. *If* $\mathbf{F}(\mathbf{A}) = \sum_{p=0}^{\infty} c_p \mathbf{A}^p$ *is an analytic matrix function and* $\mathbf{B}$ *is the block matrix in Eq.* (94), *then*

$$\mathbf{F}(\mathbf{B}) = \begin{bmatrix} \mathbf{F}(\mathbf{A}) & \mathbf{D}_{\mathbf{V}}^{[1]}(\mathbf{F}(\mathbf{A}))/1! & \cdots & \mathbf{D}_{\mathbf{V}}^{[q]}(\mathbf{F}(\mathbf{A}))/q! \\ \mathbf{0} & \mathbf{F}(\mathbf{A}) & \cdots & \mathbf{D}_{\mathbf{V}}^{[q-1]}(\mathbf{F}(\mathbf{A}))/(q-1)! \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{F}(\mathbf{A}) \end{bmatrix}. \tag{102}$$

Proof. This follows at once from Proposition 4.12 together with the fact that $\sum_{p=0}^{\infty} c_p \mathbf{B}^p$ is a block semi-circulant matrix whose blocks are sums of the form $\sum_{p=0}^{\infty} c_p \mathbf{D}_{\mathbf{V}}^{[k]}(\mathbf{A}^p)/k!$ $(k = 0, \ldots, q)$. ∎

The equations in (101) are generalizations of those in Eq. (60), and can be used in computing the first q directional derivatives of matrix-valued functions. Although this approach should be reasonably efficient (if one takes advantage of the semi-circulant property), it does not take into account the simplifications that are possible when $\|[\mathbf{V},\mathbf{A}]\|$ is small.

There exist several other series representations of $\mathbf{D}_{\mathbf{V}}(t,\mathbf{A})$, the first one due to Hausdorff [29, 42]. It takes the form of a commutator series, and is derived by applying the exponential commutator formula in Eq. (48) to the integral representation (2) as follows:

$$\begin{aligned}\mathbf{D}_{\mathbf{V}}(t,\mathbf{A}) &= e^{t\mathbf{A}}\int_0^t e^{-\tau\mathbf{A}}\mathbf{V}e^{\tau\mathbf{A}}\,d\tau = e^{t\mathbf{A}}\int_0^t\{\mathbf{V},e^{\tau\mathbf{A}}\}\,d\tau = e^{t\mathbf{A}}\left\{\mathbf{V},\int_0^t e^{\tau\mathbf{A}}\,d\tau\right\}\\ &= e^{t\mathbf{A}}\left\{\mathbf{V},\frac{e^{t\mathbf{A}}-\mathbf{I}}{\mathbf{A}}\right\} = e^{t\mathbf{A}}\sum_{k=0}^{\infty}\frac{t^{k+1}}{(k+1)!}\{\mathbf{V},\mathbf{A}^k\}.\end{aligned} \tag{103}$$

The "right" version of the series in Eq. (103) follows from the corresponding "right" version of integral representation, i.e.,

$$\mathbf{D}_{\mathbf{V}}(t,\mathbf{A}) = \left(\int_0^t e^{\tau\mathbf{A}}\mathbf{V}e^{-\tau\mathbf{A}}\,d\tau\right)e^{t\mathbf{A}} = \left\{\frac{e^{t\mathbf{A}}-\mathbf{I}}{\mathbf{A}},\mathbf{V}\right\}e^{t\mathbf{A}}. \tag{104}$$

By using the time-symmetric version of the integral representation (78) instead, we obtain the following new and more rapidly converging series:

THEOREM 4.14. *The first directional derivative admits an expansion in even commutator powers, as follows*:

$$\begin{aligned}\mathbf{D}_{\mathbf{V}}(t,\mathbf{A}) &= te^{t\mathbf{A}/2}\left\{\mathbf{V},\frac{\sinh(t\mathbf{A}/2)}{t\mathbf{A}/2}\right\}e^{t\mathbf{A}/2}\\ &= te^{t\mathbf{A}/2}\sum_{k=0}^{\infty}\frac{\{\mathbf{V},(t\mathbf{A}/2)^{2k}\}}{(2k+1)!}e^{t\mathbf{A}/2}.\end{aligned} \tag{105}$$

Proof. In the time-symmetric representation (78), the integral differs from the left-most one in Eq. (103) only by its limits. Therefore the integral average of $\mathbf{V}$ is given by

$$\begin{aligned}\frac{1}{t}\int_{-t/2}^{t/2} e^{-\tau\mathbf{A}}\mathbf{V}e^{\tau\mathbf{A}}\,d\tau &= \frac{1}{t}\int_{-t/2}^{t/2}\{\mathbf{V},e^{\tau\mathbf{A}}\}\,d\tau = \frac{1}{t}\left\{\mathbf{V},\int_{-t/2}^{t/2}e^{\tau\mathbf{A}}\,d\tau\right\}\\ &= \frac{1}{t}\left\{\mathbf{V},\frac{e^{t\mathbf{A}/2}-e^{-t\mathbf{A}/2}}{\mathbf{A}}\right\} = \left\{\mathbf{V},\frac{\sinh(t\mathbf{A}/2)}{t\mathbf{A}/2}\right\}.\end{aligned} \tag{106}$$

■

By analogy to the well-known "sinc" function, we shall refer to the function $\sinh(z)/z$ as the "sinch" function, and this series in even commutator powers as the "sinch" series.

The fact that $\mathbf{D}_{\mathbf{V}}(t,\mathbf{A})$ admits an expansion in terms of a summable commutator series makes it possible to derive a closed form representation in terms of the matrix $\mathbf{K}(\mathbf{A}) = \mathbf{A}^{\mathrm{T}} \oplus (-\mathbf{A})$ defined in Eq. (42).

THEOREM 4.15. *The first directional derivative admits the closed form representations*

$$\begin{aligned}\operatorname{col}(\mathbf{D}_{\mathbf{V}}(t,\mathbf{A})) &= (\mathbf{I} \otimes e^{t\mathbf{A}})\left(\frac{e^{t\mathbf{K}(\mathbf{A})} - \mathbf{I}}{\mathbf{K}(\mathbf{A})}\right)\operatorname{col}(\mathbf{V}) \\ &= (e^{t\mathbf{A}^{\mathrm{T}}} \otimes \mathbf{I})\left(\frac{\mathbf{I} - e^{-t\mathbf{K}(\mathbf{A})}}{\mathbf{K}(\mathbf{A})}\right)\operatorname{col}(\mathbf{V}).\end{aligned} \tag{107}$$

Proof. We prove only the first equality in Eq. (107), since the proof of the second equality is similar. Operating with "col" on both sides of Eq. (103) results in

$$\operatorname{col}(\mathbf{D}_{\mathbf{V}}(t,\mathbf{A})) = (\mathbf{I} \otimes e^{t\mathbf{A}})\operatorname{col}\left\{\mathbf{V}, \frac{e^{t\mathbf{A}} - \mathbf{I}}{\mathbf{A}}\right\}. \tag{108}$$

If we let $\mathbf{F}(\mathbf{A}) = (\exp(t\mathbf{A}) - \mathbf{I})/\mathbf{A}$, Theorem 3.2 implies that $\operatorname{col}\{\mathbf{V}, \mathbf{F}(\mathbf{A})\} = \mathbf{F}(\mathbf{K}(\mathbf{A}))\operatorname{col}(\mathbf{V})$. ∎

THEOREM 4.16. *The closed form representation of the commutator series* (105) *is given by*

$$\operatorname{col}(\mathbf{D}_{\mathbf{V}}(t,\mathbf{A})) = t(e^{t\mathbf{A}^{\mathrm{T}}/2} \otimes e^{t\mathbf{A}/2})\operatorname{sinch}(t\mathbf{K}(\mathbf{A})/2)\operatorname{col}(\mathbf{V}). \tag{109}$$

Proof. On applying the "col" operator to the first equality in Eq. (105), we obtain $\operatorname{col}(\mathbf{D}_{\mathbf{V}}(t,\mathbf{A})) = t(e^{t\mathbf{A}^{\mathrm{T}}/2} \otimes e^{t\mathbf{A}/2})\operatorname{col}\{\mathbf{V}, \operatorname{sinch}(t\mathbf{A}/2)\}$. Equation (109) now follows at once from the Theorem 3.2. ∎

5. THE COMPUTATION OF THE DIRECTIONAL DERIVATIVE

If $\mathbf{A}$ is a semi-simple matrix whose eigenvalues as well as left and right eigenvectors are known, then the directional derivative $\mathbf{D}_{\mathbf{V}}(t,\mathbf{A})$ of the matrix exponential can be computed from its spectral representations in Theorem 4.5 or Theorem 4.6 in essentially four matrix multiplications. In

practice, the complete diagonalization of a real matrix requires about 25 matrix multiplications worth of work, for a total of about 29, and in many cases this may be the most efficient method of computing the directional derivative. If the eigenvalues and eigenvectors of the matrix **A** are unavailable and difficult to compute, however, then other approaches must be tried. In some cases, these other approaches may also be faster or yield a more accurate result even when **A** can be readily diagonalized.

Two such approaches have emerged. The first is based on numerical integration of the integral representation (2), using, for example, the trapezoidal rule [38] or Simpson's rule [44]. The second is based upon a truncated series expansion, using, for example, the Hausdorff series [39] or one of the others given in Section 4 of this paper. With both these approaches, a substantial number of terms in the series, or integrand values, may be needed to obtain adequate accuracy, and there is the possibility of catastrophic cancellation. These problems are greatly alleviated by an approach analogous to that used in Sections 2.2 and 2.3 to compute the matrix exponential. The idea is to decrease the order of a truncated series approximation, or the integration range, by scaling down the argument, and then use two-sided iterative squaring to recover the derivative of the original argument. Thus these two approaches differ from each other in their choice of approximation over the reduced interval, in the size of the scale factor, and by how the overall computational process is optimized.

If the argument is scaled by a binary power 2^{-d}, and the approximations to the exponential and directional derivative of the scaled argument are denoted by $\mathbf{E}_0$ and $\mathbf{D}_0$, respectively, then two-sided iterative squaring is accomplished by the loop

$$\begin{aligned} &\textit{Loop}: \quad k = 1, \ldots, d \\ &\qquad \mathbf{D}_k := \mathbf{E}_{k-1}\mathbf{D}_{k-1} + \mathbf{D}_{k-1}\mathbf{E}_{k-1} \\ &\qquad \mathbf{E}_k := \mathbf{E}_{k-1}^2 \qquad\qquad \textit{End} \end{aligned} \tag{110}$$

Note that this includes simple iterative squaring of the exponential, and likewise tends to amplify rounding errors. An example of this may be found in Appendix A3.

In this part of the paper we present two algorithms for computing the directional derivative, which are based on the series representations derived in Section 4.3. Another approach that we have also explored uses the Romberg integration procedure [31]. This approach relies implicitly on argument halving, rather than explicit argument doubling as in two-sided iterative squaring, and consists of computing a sequence of estimates for the definite integral (2), followed by an interpolation process. We have

optimized the computation of the estimates in the first column of a Romberg array, and determined the number of halvings necessary to attain any given precision by this procedure. Nevertheless, we have found that this approach is generally slower than those described below.

5.1. *The "sinch" Algorithm for the Directional Derivative*

The decay rate of its coefficients implies that the sinch series (105) derived in Section 4.3 will yield particularly efficient approximations to the directional derivative. If we define $\mathbf{B} = t\mathbf{A}$ and $\mathbf{W} = t\mathbf{V}$, then $\mathbf{D}_{\mathbf{V}}(t, \mathbf{A})$ can be rewritten as

$$\begin{aligned}\mathbf{D}_{\mathbf{W}}(1,\mathbf{B}) &= e^{\mathbf{B}/2}\{\mathbf{W}, \operatorname{sinch}(\mathbf{B}/2)\}e^{\mathbf{B}/2} \\ &= e^{\mathbf{B}/2}\left\{\mathbf{W}, \sum_{k=0}^{\nu} \frac{(\mathbf{B}/2)^{2k}}{(2k+1)!} + \sum_{k=\nu+1}^{\infty} \frac{(\mathbf{B}/2)^{2k}}{(2k+1)!}\right\}e^{\mathbf{B}/2} \qquad (111) \\ &= e^{\mathbf{B}/2}\{\mathbf{W}, \mathbf{S}_{2\nu}(\mathbf{B}/2)\}e^{\mathbf{B}/2} + e^{\mathbf{B}/2}\{\mathbf{W}, \mathbf{T}_{2\nu}(\mathbf{B}/2)\}e^{\mathbf{B}/2}.\end{aligned}$$

We now construct an algorithm for computing $\mathbf{D}_{\mathbf{W}}(1, \mathbf{B})$ using the truncated series $\{\mathbf{W}, \mathbf{S}_{2\nu}(\mathbf{B}/2)\}$ to approximate $\mathbf{D}_{\mathbf{W}}(1, \mathbf{B})$, while using the relative norm of the remainder $\|\{\mathbf{W}, \mathbf{T}_{2\nu}(\mathbf{B}/2)\}\|/\|\mathbf{W}\|$ to bound the error. This algorithm simultaneously computes both the exponential and its directional derivative.

The four steps of the algorithm are:

1. Determination of an integer $d \geq 0$ followed by argument reduction to obtain matrices $\mathbf{C} = \mathbf{B}/2^d$ and $\mathbf{X} = \mathbf{W}/2^d$ with $\|\mathbf{C}\| \leq \beta$, where β is determined by the optimization process described below.

2. Computation of $\exp(\mathbf{C}/2)$ by the method described in Section 2.2.

3. Approximation of $\mathbf{D}_{\mathbf{X}}(1, \mathbf{C})$ by means of Eq. (111), i.e., $e^{\mathbf{C}/2}\{\mathbf{X}, \mathbf{S}_{2\nu}(\mathbf{C}/2)\}e^{\mathbf{C}/2}$, where the truncated sinch commutator series $\{\mathbf{X}, \mathbf{S}_{2\nu}(\mathbf{C}/2)\}$ of degree 2ν is computed as described in Section 3.2.

4. Recovery of the directional derivative of the original argument $\mathbf{D}_{\mathbf{W}}(1, \mathbf{B})$ together with the exponential $\exp(\mathbf{B})$ by the two-sided iterative squaring algorithm in Eq. (110).

Given a real number $\beta > 0$, the argument scaling for the directional derivative is straightforward: Let d be the integer

$$d = \begin{cases} \lceil \log_2(\|\mathbf{B}\|/\beta) \rceil & \text{if } \|\mathbf{B}\| > \beta \\ 0 & \text{otherwise,} \end{cases} \qquad (112)$$

and define the scaled matrices $\mathbf{C}$ and $\mathbf{X}$ as in step 1 above. As we shall see, the value of d needed for the directional derivative is generally larger than that needed for the exponential, so that $\exp(\mathbf{C}/2)$ can be computed without iterative squaring in step 2. The two-sided iterative squaring loop in step 4 of this algorithm needs $3d$ matrix multiplications, and hence when $\log_2(\|\mathbf{B}\|/\beta) \gg 1$ this loop becomes the most time consuming step.

The calculation of β and optimization of the overall computational process depends on the remainder $\{\mathbf{X}, \mathbf{T}_{2\nu}(\mathbf{C}/2)\}$. As shown by Eq. (106), $\{\mathbf{X}, \operatorname{sinch}(\mathbf{C}/2)\}$ may be regarded as an integral average of $\mathbf{X}$, and since its remainder is also linear in $\mathbf{X}$ we define the relative error of the approximation to be $\|\{\mathbf{X}, \mathbf{T}_{2\nu}(\mathbf{C}/2)\}\|/\|\mathbf{X}\|$. This has the advantage that the upper bound on the relative error given below is independent of $\|\mathbf{X}\|$.

PROPOSITION 5.1. *The relative error* $\|\{\mathbf{X}, \mathbf{T}_{2\nu}(\mathbf{C}/2)\}\|/\|\mathbf{X}\|$ *is bounded by*

$$\frac{\|\{\mathbf{X}, \mathbf{T}_{2\nu}(\mathbf{C}/2)\}\|}{\|\mathbf{X}\|} \leq \frac{\|\mathbf{C}\|^{2(\nu+1)}}{(2\nu+3)!}\omega_{2\nu+3}(\|\mathbf{C}\|). \tag{113}$$

where $\omega_{2\nu+3}(\alpha)$ *is the even entire function*

$$\omega_{2\nu+3}(\alpha) = 1 + (2\nu+3)! \sum_{k=1}^{\infty} \frac{\alpha^{2k}}{(2\nu+3+2k)!}. \tag{114}$$

Proof. Taking the norm of the series expansion for $\{\mathbf{X}, \mathbf{T}_{2\nu}(\mathbf{C}/2)\}$ (see Eq. (111)), and using the fact that for any commutator power $\|\{\mathbf{X}, (\mathbf{C}/2)^k\}\| \leq \|\mathbf{X}\|\,\|\mathbf{C}\|^k$, we obtain Eq. (113) after obvious algebraic simplifications. That $\omega_{2\nu+3}(\alpha)$ is an entire function follows easily from the ratio test. ∎

The optimal combination of the degree ν and the norm $\alpha = \|\mathbf{C}\|$ can now be determined. Given any desired accuracy $\varepsilon > 0$, we equate it to the bound in the inequality (113), which gives rise to the sequence of nonlinear equations

$$\frac{\alpha^{2(\nu+1)}}{(2\nu+3)!}\omega_{2\nu+3}(\alpha) = \varepsilon \tag{115}$$

for α_ν $(\nu = 0, 1, \ldots)$. Since for each ν, the function $\omega_{2\nu+3}(\alpha) \geq 1$ is convex as well as monotonically increasing for $\alpha \geq 0$, it follows that the nonlinear equation (115) possesses a unique positive solution α_ν. Furthermore, Newton's method applied to Eq. (115) generates a monotonically decreasing sequence of iterates $\alpha_\nu^{[i]}$ that converges quadratically to this solution provided that it starts from the overestimate

$$\alpha_\nu^{[0]} = \left(\varepsilon(2\nu+3)!\right)^{1/2(\nu+1)}. \tag{116}$$

Thus for any given ε we can easily construct a table (ν, α_ν), which is used in the optimization procedure that we now describe.

The optimality criterion for choosing a particular combination of α_ν and ν from this table is the minimization of the total number of matrix multiplications $\mu(\nu)$ needed by the algorithm. If we assume that the work needed to approximate the reduced exponential $\exp(\mathbf{C}/2)$ varies slowly with α_ν and may be bounded by a constant η, then step 3 requires $4\nu + 2$ matrix multiplications for computing $\mathbf{D}_0 = \mathbf{D}_\mathbf{X}(1, \mathbf{C})$ and $\eta + 1$ for the rest. The iterative squaring loop of step 4, on the other hand, requires three matrix multiplications to double the interval, two for the directional derivative, and one for the matrix exponential. Therefore, the function to be minimized is given by

$$\mu(\nu) = 4\nu + \eta + 3 + 3\lceil d(\alpha_\nu)\rceil = 4\nu + \eta + 3 + 3\lceil \log_2(\|\mathbf{B}\|/\alpha_\nu)\rceil. \tag{117}$$

It can be shown that the sequence of real numbers $\{\log_2(\alpha_\nu)\}$ is a monotonic sequence which grows asymptotically as $\log_2(2\nu + 3)$. Because the coefficients 3 and 4 are relatively prime, the 4ν term cannot be moved inside the ceiling function in Eq. (117). This can cause the value of the minimum to fluctuate by one, depending on the fractional value of $\|\mathbf{B}\|$. Nevertheless, for a given value of $\|\mathbf{B}\|$ this minimum occurs at some integer ν^*, which determines the scale factor $\beta = \alpha_{\nu^*}$ to be used in Eq. (112).

Due to this fluctuation, in single precision $\varepsilon = 2^{-23}$ the optimal combination is either $\nu^* = 2$ or $\nu^* = 3$ with a corresponding value of $\beta = 0.290388$ or $\beta = 0.674786$. The *a posteriori* check using Table I in Section 2.3 shows that $\eta < 4$. Similarly, for double precision $\varepsilon = 2^{-52}$ the optimal combination is either $\nu^* = 3$ or $\nu^* = 4$ with a corresponding value of $\beta = 0.054736$ or $\beta = 0.156585$, and $\eta < 5$. It follows by a short calculation that in single precision, spectral methods will generally be more efficient whenever more than about five iterative squarings are required, i.e., $\|\mathbf{B}\| > 2^5 \times 0.290388 \approx 9.29$, or three iterative squarings in double precision, i.e., $\|\mathbf{B}\| > 2^3 \times 0.054736 \approx 0.44$.

5.2. *The "exp" Algorithm for the Directional Derivative*

This algorithm for computing $\mathbf{D}_\mathbf{W}(\mathbf{B})$ is based on the functional representation given in Proposition 4.10, namely

$$e^{2\mathbf{Z}} = \exp\left(2\begin{bmatrix}\mathbf{B} & \mathbf{W}\\ \mathbf{0} & \mathbf{B}\end{bmatrix}\right) = \begin{bmatrix}\exp(2\mathbf{B}) & \mathbf{D}_\mathbf{W}(2, \mathbf{B})\\ \mathbf{0} & \exp(2\mathbf{B})\end{bmatrix} \tag{118}$$

together with a slight modification of the algorithm for computing the matrix exponential described in Sections 2.2 and 2.3.

By applying the rational approximation of the exponential derived in Section 2.2 to exp(2**Z**), we immediately obtain a rational approximation to the directional derivative. Due to the block triangular structure of the matrix argument **Z**, this approximation can be calculated in substantially less time than that required to approximate the exponential of an arbitrary $2n \times 2n$ matrix. Using Eqs. (26) and (33) we may write

$$e^{2\mathbf{Z}} = \frac{\mathbf{H}(\mathbf{Z}) + \mathbf{Z}}{\mathbf{H}(\mathbf{Z}) - \mathbf{Z}} = \frac{\mathbf{M}(\mathbf{Z}) + \mathbf{Z}\mathbf{N}(\mathbf{Z})}{\mathbf{M}(\mathbf{Z}) - \mathbf{Z}\mathbf{N}(\mathbf{Z})}$$

$$= \frac{\begin{bmatrix} \mathbf{M} & \mathbf{P} \\ \mathbf{0} & \mathbf{M} \end{bmatrix} + \begin{bmatrix} \mathbf{B} & \mathbf{W} \\ \mathbf{0} & \mathbf{B} \end{bmatrix}\begin{bmatrix} \mathbf{N} & \mathbf{Q} \\ \mathbf{0} & \mathbf{N} \end{bmatrix}}{\begin{bmatrix} \mathbf{M} & \mathbf{P} \\ \mathbf{0} & \mathbf{M} \end{bmatrix} - \begin{bmatrix} \mathbf{B} & \mathbf{W} \\ \mathbf{0} & \mathbf{B} \end{bmatrix}\begin{bmatrix} \mathbf{N} & \mathbf{Q} \\ \mathbf{0} & \mathbf{N} \end{bmatrix}}. \tag{119}$$

where $\mathbf{M} = \mathbf{M}(\mathbf{B})$ and $\mathbf{N} = \mathbf{N}(\mathbf{B})$ designate the numerator and denominator of the rational representation of $\mathbf{H}(\mathbf{B}) = \cosh(\mathbf{B})/\sinh(\mathbf{B})$, and

$$\mathbf{P} = \alpha_2 \mathbf{D}_{\mathbf{W}}(\mathbf{B}^2) + \alpha_4 \mathbf{D}_{\mathbf{W}}(\mathbf{B}^4) + \cdots$$

$$\mathbf{Q} = \beta_2 \mathbf{D}_{\mathbf{W}}(\mathbf{B}^2) + \beta_4 \mathbf{D}_{\mathbf{W}}(\mathbf{B}^4) + \cdots . \tag{120}$$

Here, α_{2k} and β_{2k} denote the coefficients of $\mathbf{B}^{2k}$ in **M** and **N**, respectively, $\mathbf{D}_{\mathbf{W}}(\mathbf{B}^{2k})$ is the $2k$th power directional derivative, and hence **P** and **Q** are the directional derivatives of cosh(**B**) and sinch(**B**) in the direction **W**.

If we let $\mathbf{T} = \mathbf{BQ} + \mathbf{WN}$ and observe that the inverse of any block triangular matrix is given by

$$\begin{bmatrix} \mathbf{K} & \mathbf{L} \\ \mathbf{0} & \mathbf{K} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{K}^{-1} & -\mathbf{K}^{-1}\mathbf{L}\mathbf{K}^{-1} \\ \mathbf{0} & \mathbf{K}^{-1} \end{bmatrix} \tag{121}$$

whenever **K** is invertible, then Eq. (119) can be written as

$$e^{2\mathbf{Z}} = \frac{\begin{bmatrix} \mathbf{M} + \mathbf{BN} & \mathbf{P} + \mathbf{T} \\ \mathbf{0} & \mathbf{M} + \mathbf{BN} \end{bmatrix}}{\begin{bmatrix} \mathbf{M} - \mathbf{BN} & \mathbf{P} - \mathbf{T} \\ \mathbf{0} & \mathbf{M} - \mathbf{BN} \end{bmatrix}} = \begin{bmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{0} & \mathbf{E} \end{bmatrix}. \tag{122}$$

The matrix $\mathbf{E} = (\mathbf{M} - \mathbf{BN})^{-1}(\mathbf{M} + \mathbf{BN})$ is the same rational representation of exp(2**B**) as that in Eq. (33), while $\mathbf{F} = (\mathbf{M} - \mathbf{BN})^{-1}(\mathbf{P} + \mathbf{T} - (\mathbf{P} - \mathbf{T})\mathbf{E})$ is a rational representation of the directional derivative $\mathbf{D}_{2\mathbf{W}}(2\mathbf{B})$.

In this case, the four stages of the algorithm are:

1. Determination of an integer $d \geq 0$ followed by argument reduction to obtain matrices $\mathbf{C} = \mathbf{B}/2^{d+1}$ and $\mathbf{X} = \mathbf{W}/2^{d+1}$ satisfying

$$\left\| \begin{bmatrix} \mathbf{C} & \mathbf{X} \\ \mathbf{0} & \mathbf{C} \end{bmatrix}^2 \right\| \leq \gamma. \tag{123}$$

where γ is obtained from Table I in Section 2.3.

2. Calculation of the numerator and denominator polynomials $\mathbf{M}_{2m}(\mathbf{C})$ and $\mathbf{N}_{2m}(\mathbf{C})$ using the Padé table (32) and, using these same coefficients, formation of the matrices

$$\begin{aligned} \mathbf{P}_{2m} &= \alpha_2^{2m}\mathbf{D}_{\mathbf{X}}(\mathbf{C}^2) + \alpha_4^{2m}\mathbf{D}_{\mathbf{X}}(\mathbf{C}^4) + \cdots + \alpha_{2m}^{2m}\mathbf{D}_{\mathbf{X}}(\mathbf{C}^{2m}) \\ \mathbf{Q}_{2m} &= \beta_2^{2m}\mathbf{D}_{\mathbf{X}}(\mathbf{C}^2) + \beta_4^{2m}\mathbf{D}_{\mathbf{X}}(\mathbf{C}^4) + \cdots + \beta_{2m}^{2m}\mathbf{D}_{\mathbf{X}}(\mathbf{C}^{2m}) \\ \mathbf{T} &= \mathbf{C}\mathbf{Q}_{2m} + \mathbf{X}\mathbf{N}_{2m} \\ \mathbf{G} &= \mathbf{C}\mathbf{N}_{2m}. \end{aligned} \tag{124}$$

3. Computation of the LU decomposition of $\mathbf{M}_{2m} - \mathbf{G}$ to be used in the left-matrix divisions

$$\begin{aligned} \mathbf{E}_0 &= (\mathbf{M}_{2m} - \mathbf{G})^{-1}(\mathbf{M}_{2m} + \mathbf{G}) \\ \mathbf{D}_0 &= (\mathbf{M}_{2m} - \mathbf{G})^{-1}(\mathbf{P}_{2m} + \mathbf{T} - (\mathbf{P}_{2m} - \mathbf{T})\mathbf{E}_0), \end{aligned} \tag{125}$$

which yield the requisite approximations to $\exp(2\mathbf{C})$ and $\mathbf{D}_{\mathbf{X}}(2, \mathbf{C})$.

4. Recovery of the directional derivative of the original argument $\mathbf{D}_{\mathbf{W}}(1, \mathbf{B})$ and the exponential $\exp(\mathbf{B})$ by d two-sided iterative squarings.

As before, the optimization consists of minimizing the total number $\zeta(m)$ of matrix multiplications for this process. Using Table I in Section 2.3, which gives γ_{2m} as a function of m, a count of the multiplications in steps 2, 3, and 4 shows that $\zeta(m)$ can be written as

$$\begin{aligned} \zeta(m) &= 6\tfrac{2}{3} + 3m + 3d(\gamma_{2m}) = 6\tfrac{2}{3} + 3m + 3\left\lceil \tfrac{1}{2}\log_2\left(\|\mathbf{Z}^2\|/4\gamma_{2m}\right)\right\rceil \\ &= 3\tfrac{2}{3} + 3\left\lceil m - \tfrac{1}{2}\log_2(\gamma_{2m}) + \tfrac{1}{2}\log_2(\|\mathbf{Z}^2\|)\right\rceil. \end{aligned} \tag{126}$$

Comparison of this last expression with that in Eq. (38) shows that $\zeta(m) = \zeta_0 + 3\eta(m)$ for some constant ζ_0, where $\eta(m)$ is the work required to compute $\exp(\mathbf{B})$ in units of matrix multiplications. It follows that the minima of $\zeta(m)$ and $\eta(m)$ occur at the same integer m^*, with the same associated value of $\gamma = \gamma_{2m^*}$ from Table I. This gives a total amount

of work equivalent to

$$\zeta(m^*) = \tfrac{11}{3} + 3\left[\kappa(\varepsilon) + \tfrac{1}{2}\log_2(\|\mathbf{Z}^2\|)\right] \tag{127}$$

matrix multiplications, where $\kappa(\varepsilon)$ is given by Eq. (39) in Section 2.3.

Unlike the "sinch" algorithm, the "exp" algorithm allows one to use the partial spectral product decomposition in order to remove the few largest eigenvalues, which contribute most to the rounding error propagation in two-sided iterative squaring.

6. THE GRADIENT OF THE MATRIX EXPONENTIAL

For simplicity, in this part of the paper we restrict our attention to real square matrices only. If $\mathbf{E}_{kl}$ denotes the $[k,l]$th elementary matrix (with a one in the $[k,l]$th position and zeros elsewhere) then $\mathbf{D}_{\mathbf{E}_{kl}}(1,\mathbf{A}) = \mathbf{D}_{\mathbf{E}_{kl}}(\exp(\mathbf{A}))$ is called the $[k,l]$th directional derivative. It measures the effect of a perturbation in a single entry a_{kl} of the matrix $\mathbf{A}$ on the entire matrix expotential, and will be written as

$$\frac{\partial e^{\mathbf{A}}}{\partial a_{kl}} = \mathbf{D}_{\mathbf{E}_{kl}}(1,\mathbf{A}). \tag{128}$$

On the other hand, each entry of $\exp(\mathbf{A})$ is a scalar function of all n^2 entries of $\mathbf{A}$, and therefore it makes sense to arrange the gradient of the $[k,l]$th entry of the matrix exponential in the form of a matrix [27]

$$\frac{\partial e^{\mathbf{A}}_{kl}}{\partial \mathbf{A}} = \begin{bmatrix} \partial e^{\mathbf{A}}_{kl}/\partial a_{11} & \cdots & \partial e^{\mathbf{A}}_{kl}/\partial a_{1n} \\ \cdots & \cdots & \cdots \\ \partial e^{\mathbf{A}}_{kl}/\partial a_{n1} & \cdots & \partial e^{\mathbf{A}}_{kl}/\partial a_{nn} \end{bmatrix}. \tag{129}$$

We shall now determine the relation between the matrices $\partial e^{\mathbf{A}}/\partial a_{kl}$ and $\partial e^{\mathbf{A}}_{kl}/\partial \mathbf{A}$. First, we need a result relating the $[k,l]$th directional derivative of the power function $\mathbf{Y} = \mathbf{Y}(\mathbf{A}) = \mathbf{A}^p$ $(p \geq 1)$ to the gradient of $y_{kl}(\mathbf{A})$, the $[k,l]$th entry of the matrix $\mathbf{Y}(\mathbf{A})$.

PROPOSITION 6.1. *If the entries of* $\mathbf{A}$ *are independent, then*

$$\frac{\partial y_{kl}(\mathbf{A})}{\partial \mathbf{A}} = \frac{\partial \mathbf{Y}(\mathbf{A}^{\mathsf{T}})}{\partial a_{kl}} = \left[\frac{\partial \mathbf{Y}(\mathbf{A})}{\partial a_{lk}}\right]^{\mathsf{T}}. \tag{130}$$

Proof. The directional derivative of $\mathbf{Y}(\mathbf{A})$ in the $[k,l]$th direction is given by Eq. (5) with $\mathbf{V} = \mathbf{E}_{kl}$:

$$\frac{\partial \mathbf{Y}(\mathbf{A})}{\partial a_{kl}} = \mathbf{D}_{\mathbf{E}_{kl}}(\mathbf{A}^p) = \sum_{i=1}^{p} \mathbf{A}^{p-i}\mathbf{E}_{kl}\mathbf{A}^{i-1}. \tag{131}$$

The $[k,l]$th gradient of $\mathbf{Y}$, on the other hand, is given by (see, for example, [27])

$$\frac{\partial y_{kl}(\mathbf{A})}{\partial \mathbf{A}} = \sum_{i=1}^{p} (\mathbf{A}^{\mathrm{T}})^{p-i}\mathbf{E}_{kl}(\mathbf{A}^{\mathrm{T}})^{i-1}, \tag{132}$$

which, on comparison with Eq. (131), proves the first equality in Eq. (130). To prove the second, we transpose both sides of Eq. (132) to obtain

$$\begin{aligned}\left[\frac{\partial y_{kl}(\mathbf{A})}{\partial \mathbf{A}}\right]^{\mathrm{T}} &= \sum_{i=1}^{p} \mathbf{A}^{i-1}\mathbf{E}_{kl}^{\mathrm{T}}\mathbf{A}^{p-i} \\ &= \sum_{i=1}^{p} \mathbf{A}^{i-1}\mathbf{E}_{lk}\mathbf{A}^{p-i} = \frac{\partial \mathbf{Y}(\mathbf{A})}{\partial a_{lk}},\end{aligned} \tag{133}$$

where the order of summation is reversed in the last sum. ∎

We now specialize this result to symmetric matrices (whose entries are not independent). Let $\mathbf{Z}(\mathbf{B})$ denote the restriction of $\mathbf{Y}(\mathbf{A})$ to the set of symmetric matrices. Then the entries $z_{kl}(\mathbf{B})$ and $z_{lk}(\mathbf{B})$ denote the same function, since any power of a symmetric matrix is symmetric. Hence, the gradient matrices $\partial z_{kl}/\partial \mathbf{B}$ and $\partial z_{lk}/\partial \mathbf{B}$ are also the same. Furthermore, this matrix is symmetric, since by the definition of the gradient [27] $(\partial z_{kl}/\partial \mathbf{B})^{\mathrm{T}} = \partial z_{kl}/\partial(\mathbf{B}^{\mathrm{T}}) = \partial z_{kl}/\partial \mathbf{B}$. In the following, we shall need the symmetrization operator

$$\mathrm{sym}(\mathbf{A}) = \mathbf{A} + \mathbf{A}^{\mathrm{T}} - \mathrm{diag}(\mathbf{A}), \tag{134}$$

where $\mathrm{diag}(\mathbf{A})$ is the diagonal matrix whose elements are a_{kk} $(k = 1,\ldots,n)$. We note that the sym operator also appears in [8] in connection with differentiating the trace of various matrix functions. This operator is of much broader interest, however, as shown below.

PROPOSITION 6.2. *If* $\mathbf{B}$ *is a symmetric matrix, the gradient of* $z_{kl}(\mathbf{B})$ *with respect to* $\mathbf{B}$ *is*

$$\frac{\partial z_{kl}(\mathbf{B})}{\partial \mathbf{B}} = \mathrm{sym}\left(\frac{\partial \mathbf{Y}(\mathbf{A})}{\partial a_{kl}}\bigg|_{\mathbf{A}=\mathbf{B}}\right), \tag{135}$$

where the directional derivative $\partial \mathbf{Y}(\mathbf{A})/\partial a_{kl}$ *in the nonsymmetric direction* $\mathbf{E}_{kl}$ *is evaluated at the symmetric matrix* $\mathbf{B}$.

Proof. For any $i \neq j$, consider $y_{kl}(\mathbf{A})$ to be a function of two variables $y_{kl}(a_{ij}, a_{ji})$ only (with the remaining variables held fixed). Similarly, let $z_{kl}(b_{ij})$ denote the function of one variable that is obtained from $z_{kl}(\mathbf{B})$ by holding all variables except $b_{ij} = b_{ji}$ fixed. Let $\iota\colon \mathbf{R} \to \mathbf{R}^2$ be the mapping $b_{ij} \to [b_{ij}, b_{ij}]$. Then $z_{kl}(b_{ij}) = y_{kl}(\iota(b_{ij}))$, so that by the chain rule we have

$$\begin{aligned} \frac{\partial z_{kl}(\mathbf{B})}{\partial b_{ij}} &= \frac{\partial z_{kl}(b_{ij})}{\partial b_{ij}} = \nabla y_{kl}\big(\iota(b_{ij})\big) \cdot \frac{d\iota(b_{ij})}{db_{ij}} \\ &= \left[\frac{\partial y_{kl}}{\partial a_{ij}}, \frac{\partial y_{kl}}{\partial a_{ji}}\right] \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{\partial y_{kl}}{\partial a_{ij}} + \frac{\partial y_{kl}}{\partial a_{ji}}\bigg|_{a_{ij} = a_{ji} = b_{ij}} . \end{aligned} \tag{136}$$

When $i = j$, on the other hand, there is only one variable a_{ii}, so that

$$\frac{\partial z_{kl}(\mathbf{B})}{\partial b_{ii}} = \frac{\partial y_{kl}(\mathbf{A})}{\partial a_{ii}}\bigg|_{\mathbf{A} = \mathbf{B}} . \tag{137}$$

It follows from Eq. (136) that the off-diagonal entries of $\partial z_{kl}/\partial \mathbf{B}$ are the same as those in

$$\frac{\partial y_{kl}(\mathbf{A})}{\partial \mathbf{A}} + \left(\frac{\partial y_{kl}(\mathbf{A})}{\partial \mathbf{A}}\right)^{\mathrm{T}}\bigg|_{\mathbf{A} = \mathbf{B}}, \tag{138}$$

whereas by Eq. (137) the diagonal elements are the same as those in $\partial y_{kl}/\partial \mathbf{A}$ evaluated at $\mathbf{A} = \mathbf{B}$, thus showing that

$$\frac{\partial z_{kl}(\mathbf{B})}{\partial \mathbf{B}} = \operatorname{sym}\left(\frac{\partial y_{kl}(\mathbf{A})}{\partial \mathbf{A}}\bigg|_{\mathbf{A} = \mathbf{B}}\right). \tag{139}$$

By Proposition 6.1, $\partial \mathbf{Y}(\mathbf{A})/\partial a_{kl} = \partial y_{kl}(\mathbf{A})/\partial \mathbf{A}$ when $\mathbf{A} = \mathbf{B}$ is symmetric, so the right-hand side of Eq. (139) can be replaced by the corresponding directional derivatives. ∎

Since the directional derivatives of a symmetric matrix in the elementary symmetric directions, $\hat{\mathbf{E}}_{kl} = \operatorname{sym}(\mathbf{E}_{kl})$, are often encountered in applications it is useful to relate them to gradients of a symmetric matrix. In what follows δ_{kl} denotes the Kronecker delta.

PROPOSITION 6.3. *If* $\partial \mathbf{Z}(\mathbf{B})/\partial b_{kl}$ *denotes the directional derivative in the elementary symmetric direction* $\hat{\mathbf{E}}_{kl}$, *then*

$$\frac{\partial z_{kl}(\mathbf{B})}{\partial \mathbf{B}} = (1 + \delta_{kl})\left[\frac{\partial \mathbf{Z}(\mathbf{B})}{\partial b_{kl}} - \frac{1}{2}\operatorname{diag}\left(\frac{\partial \mathbf{Z}(\mathbf{B})}{\partial b_{kl}}\right)\right]. \tag{140}$$

Proof. The case $k = l$ follows from Eq. (135), since $\partial \mathbf{Z}(\mathbf{B})/\partial b_{kk} = \partial \mathbf{Y}/\partial a_{kk}|_{\mathbf{A}=\mathbf{B}}$. For $k \neq l$ the definition of the elementary symmetric matrices and the linearity of the (nonsymmetric) directional derivative with respect to its direction allow us to write $\partial \mathbf{Z}(\mathbf{B})/\partial b_{kl} = (\partial \mathbf{Y}/\partial a_{kl} + \partial \mathbf{Y}/\partial a_{lk})|_{\mathbf{A}=\mathbf{B}}$. Since $\mathbf{A} = \mathbf{B}$ is a symmetric matrix, we have $\partial \mathbf{Y}/\partial a_{kl}|_{\mathbf{A}=\mathbf{B}} = (\partial \mathbf{Y}/\partial a_{lk}|_{\mathbf{A}=\mathbf{B}})^{\mathrm{T}}$ and hence for arbitrary indices k, l,

$$(1 + \delta_{kl})\,\partial \mathbf{Z}(\mathbf{B})/\partial b_{kl} = \partial \mathbf{Y}/\partial a_{kl} + (\partial \mathbf{Y}/\partial a_{kl})^{\mathrm{T}}\big|_{\mathbf{A}=\mathbf{B}}. \tag{141}$$

Making this substitution in Eq. (135) and observing that $\operatorname{diag}(\mathbf{A}) = \operatorname{diag}(\mathbf{A} + \mathbf{A}^{\mathrm{T}})/2$ now lead at once to the desired result. ∎

We shall denote the directional derivative of an analytic function of a matrix argument $\mathbf{F}(\mathbf{A})$ in the $[k, l]$th direction by $\partial \mathbf{F}(\mathbf{A})/\partial a_{kl}$, and the gradient of its $[k, l]$th entry by $\partial f_{kl}(\mathbf{A})/\partial \mathbf{A}$. By a slight abuse of notation, we shall use these same symbols for the function of $n(n + 1)/2$ variables obtained when $\mathbf{A}$ is replaced by a symmetric matrix $\mathbf{B}$. Then Propositions 6.1 through 6.3 can be extended in a straightforward way to arbitrary analytic functions, which makes the computation of their gradients and directional derivatives essentially equivalent.

THEOREM 6.4. *If $\mathbf{F}(\mathbf{A})$ is an analytic function of a matrix $\mathbf{A}$ whose entries are independent and such that the spectral radius of $\mathbf{A}$ is within the radius of convergence of $\mathbf{F}$, then the gradient and directional derivatives satisfy*

$$\partial f_{kl}(\mathbf{A})/\partial \mathbf{A} = \partial \mathbf{F}(\mathbf{A}^{\mathrm{T}})/\partial a_{kl} = (\partial \mathbf{F}(\mathbf{A})/\partial a_{lk})^{\mathrm{T}} \qquad (k, l = 1, \ldots, n). \tag{142}$$

In the case of a symmetric matrix $\mathbf{B}$

$$\begin{aligned} \partial f_{kl}(\mathbf{B})/\partial \mathbf{B} &= \operatorname{sym}(\partial \mathbf{F}(\mathbf{A})/\partial a_{kl})\big|_{\mathbf{A}=\mathbf{B}} \qquad (k, l = 1, \ldots, n) \\ &= (1 + \delta_{kl})\big(\partial \mathbf{F}(\mathbf{B})/\partial b_{kl} - \tfrac{1}{2}\operatorname{diag}(\partial \mathbf{F}(\mathbf{B})/\partial b_{kl})\big). \end{aligned} \tag{143}$$

Proof. By the linearity of the matrix gradient operator and Proposition 6.1, we obtain

$$\begin{aligned} \frac{\partial f_{kl}(\mathbf{A})}{\partial \mathbf{A}} &= \frac{\partial}{\partial \mathbf{A}}\left(\sum_{p=0}^{\infty} c_p \mathbf{A}^p\right)_{kl} = \sum_{p=0}^{\infty} c_p \frac{\partial}{\partial \mathbf{A}}(\mathbf{A}^p)_{kl} = \sum_{p=0}^{\infty} c_p \frac{\partial\left((\mathbf{A}^{\mathrm{T}})^p\right)}{\partial a_{kl}} \\ &= \frac{\partial \mathbf{F}(\mathbf{A}^{\mathrm{T}})}{\partial a_{kl}} = \sum_{p=0}^{\infty} c_p \left(\frac{\partial(\mathbf{A}^p)}{\partial a_{lk}}\right)^{\mathrm{T}} = \left(\frac{\partial \mathbf{F}(\mathbf{A})}{\partial a_{lk}}\right)^{\mathrm{T}}, \end{aligned} \tag{144}$$

where we have also used the linearity of the directional derivative in its direction. The proof of Eq. (143) follows similarly from Propositions 6.2 and 6.3. ∎

As a consequence, the gradients and directional derivatives of $\exp(t\mathbf{A})$ are related by (see Proposition 4.1(a))

$$\frac{\partial e^{t\mathbf{A}}_{kl}}{\partial \mathbf{A}} = t\frac{\partial e^{t\mathbf{A}}_{kl}}{\partial (t\mathbf{A})} = t\mathbf{D}_{\mathbf{E}_{kl}}(1, t\mathbf{A}^{\mathrm{T}}) = \mathbf{D}_{\mathbf{E}_{kl}}(t, \mathbf{A}^{\mathrm{T}}). \tag{145}$$

whereas for symmetric matrices $\mathbf{B}$, the corresponding symmetric gradient $\partial(\exp(t\mathbf{B}))_{kl}/\partial\mathbf{B}$ is given by

$$\frac{\partial e^{t\mathbf{B}}_{kl}}{\partial \mathbf{B}} = \operatorname{sym}\left(\frac{\partial e^{t\mathbf{A}^{\mathrm{T}}}}{\partial a_{kl}}\bigg|_{\mathbf{A}=\mathbf{B}} \right) = \operatorname{sym}\big(\mathbf{D}_{\mathbf{E}_{kl}}(t, \mathbf{B})\big). \tag{146}$$

If, on the other hand, $\partial e^{t\mathbf{B}}/\partial b_{kl} = \mathbf{D}_{\hat{\mathbf{E}}_{kl}}(t, \mathbf{B})$ denotes the directional derivative in the elementary symmetric direction, then

$$\begin{aligned} \frac{\partial e^{t\mathbf{B}}_{kl}}{\partial \mathbf{B}} &= (1 + \delta_{kl})\left(\frac{\partial e^{t\mathbf{B}}}{\partial b_{kl}} - \frac{1}{2}\operatorname{diag}\left(\frac{\partial e^{t\mathbf{B}}}{\partial b_{kl}} \right) \right) \\ &= (1 + \delta_{kl})\left(\mathbf{D}_{\hat{\mathbf{E}}_{kl}}(t, \mathbf{B}) - \frac{1}{2}\operatorname{diag}\big(\mathbf{D}_{\hat{\mathbf{E}}_{kl}}(t, \mathbf{B})\big) \right). \end{aligned} \tag{147}$$

7. AN APPLICATION TO NUCLEAR MAGNETIC RESONANCE

An important recent development in molecular biophysics, which provided the motivation for much of this paper, is the determination of protein structure from Nuclear Magnetic Resonance (NMR) spectroscopy (see, e.g., [71]). In this technique, an aqueous solution of the protein is placed in a strong, uniform magnetic field, which causes the magnetic dipoles of the hydrogen nuclei in the protein to partially align with the field. This alignment is disrupted by a brief pulse of radio-frequency radiation, and the decay of the magnetizations of the individual nuclei back to their equilibrium alignment is monitored by measuring the radiation they emit at characteristic frequencies. By selectively irradiating one of these resonances, these experiments enable the transfer of magnetization or *relaxation* of the corresponding nucleus by another, nearby nucleus

to be measured. This is known as the Nuclear Overhauser Effect (NOE).

The magnitudes of the NOEs are more easily measured by means of a sophisticated NMR experiment known as NOESY [20]. This yields a *two-dimensional* NMR spectrum, which is a real-valued bivariate function $S(\omega_1, \omega_2)$ of the frequencies of the hydrogen nuclei whose "cross-peak" intensities m_{kl} are proportional to the NOE between the nuclei. The spectrum also depends implicitly on a certain time constant in the experiment, which is known as the *mixing time* τ. The dependence of the matrix $\mathbf{M}(\tau) = [m_{kl}]$ on the mixing time is governed by the *Solomon equations* [20]

$$\frac{d\mathbf{M}}{d\tau} = -\mathbf{R} \cdot \mathbf{M}, \qquad \mathbf{M}(0) = \mathbf{I}, \tag{148}$$

where $\mathbf{R}$ is an $N \times N$ symmetric matrix known as the *relaxation matrix*. Thus the NOE intensity matrix $\mathbf{M}(\tau)$ is given by the matrix exponential $\exp(-\tau\mathbf{R})$. The NOESY spectrum itself is then obtained from a certain matrix transformation of $\mathbf{M}(\tau)$ [30].

The basic theory of NMR implies that $\mathbf{R}$ is a diagonally dominant matrix, and that for large molecules such as proteins the off-diagonal relaxation rates are nonpositive. Thus $\mathbf{R}$ is actually an M-matrix, and hence $\exp(-\tau\mathbf{R})$ is a nonnegative matrix (see a proof in Ref. [6]). Since the entries of $\mathbf{R}$ (called relaxation rates) contain information about the distances among hydrogen nuclei that can be used to determine the spatial structure of the protein [66, 71], it is important to have a method by which $\mathbf{R}$ can be determined from the entries of $\exp(-\tau\mathbf{R})$. Unfortunately, the entries of $\exp(-\tau\mathbf{R})$ can only be indirectly inferred from the spectra, and due to problems with, e.g., peak overlap and signal-to-noise limitations no information at all is available for the majority of the entries of $\exp(-\tau\mathbf{R})$. In the following, we shall absorb the sign of $-\tau\mathbf{R}$ into $\mathbf{R}$, thus converting it into a negative-definite (stable) matrix.

The problem of determining such an $\mathbf{R}$ from an incomplete and inexact list of the entries of $\exp(\tau\mathbf{R})$, generally at multiple mixing times τ, has been approached in many ways (as reviewed in Ref. [7]). The simplest and most direct approach, first considered in [39], is to use a least-squares criterion to identify the optimum choice of relaxation rates, e.g.,

$$\min_{\mathbf{R}} \left(\chi^2(\mathbf{R}) \right) = \min_{\mathbf{R}} \left(\frac{1}{2} \sum_{i=1}^{M} \sum_{(k,l) \in \Upsilon} \left(e_{kl}^{\tau_i \mathbf{R}} - m_{kl}^{[i]} \right)^2 \right). \tag{149}$$

Here $\tau_1 < \tau_2 < \cdots < \tau_M$ is a given set of mixing times at which the measurement matrices $\mathbf{M}(\tau_i) = M^{[i]} = [m_{kl}^{[i]}]$ are obtained from NOESY

spectra. The set Υ of index pairs $k \leq l$ at which measurements are available, which can be assumed to be the same for each τ_i, completes the specification of problem.

Unfortunately, interpolating noisy data by sums of scalar exponentials with unknown decay rates is a notoriously ill-conditioned problem [28, 40], which indicates that the problem of estimating $\mathbf{R}$ by minimizing χ^2 may also be ill-conditioned. It should therefore be stressed that the above problem formulation has only recently begun to be tested, and because of the incompleteness, noise, and unknown systematic errors in the measurements the problem of quantitatively estimating all the relaxation rates from actual NMR spectra is a formidable one. The formulation given above is therefore intended primarily to illustrate how the mathematical results obtained in this paper are likely to contribute to solving such parameter estimation problems, rather than as a prescription for solving them.

7.1. *Evaluating the Gradient as a Sum of M Directional Derivatives*

Given that $\mathbf{R}$ is typically of order $N = 500$ or more and that the number of independent entries in $\mathbf{R}$ is $N(N + 1)/2$, such memory intensive minimization algorithms as the Gauss–Newton or variable metric algorithms are not applicable to $\chi^2(\mathbf{R})$. At the same time, non-gradient algorithms such as the simplex method are not efficient enough to solve problems of this size. A reasonable compromise is to use first derivative algorithms such as conjugate gradients or the Levenberg–Marquardt algorithm. In principle, the number of variables can be reduced if some *a priori* knowledge of the sparsity of $\mathbf{R}$ is available, as, for example, if an approximate structure is already known for the protein. In the following, however, we shall assume we are dealing with a full matrix.

With any first derivative method, it is essential to have an efficient algorithm that evaluates the function $\chi^2(\mathbf{R})$ together with its gradient $\nabla\chi^2(\mathbf{R})$. In this section, we present some results which enable us to use the spectral, series approximation or any other convenient method of calculating directional derivatives to compute the gradient $\nabla\chi^2(\mathbf{R})$, with a complexity roughly equal to that needed to calculate the directional derivatives. These results make use of the integral representation of the directional derivative, together with the relation between the directional derivative and the gradients of the exponential's entries given in Section 6. In the next section, we will also show how the addition formulae for the exponential and its directional derivative can be used to obtain these quantities at any arithmetic or geometric sequence of mixing times, given only their values at the smallest mixing time τ_1.

To begin, we expand the gradient of $\chi^2(\mathbf{R})$ as

$$\begin{aligned}\nabla\chi^2(\mathbf{R}) &= \sum_{i=1}^{M} \sum_{(k,l)\in\Upsilon} \left(e_{kl}^{\tau_i\mathbf{R}} - m_{kl}^{[i]}\right) \frac{\partial e_{kl}^{\tau_i\mathbf{R}}}{\partial \mathbf{R}} \\ &= \sum_{i=1}^{M} \sum_{(k,l)\in\Upsilon} \xi_{kl}^{[i]} \mathbf{G}_{kl}^{[i]} = \sum_{i=1}^{M} \Gamma^{[i]},\end{aligned} \tag{150}$$

where

$$\xi_{kl}^{[i]} = e_{kl}^{\tau_i\mathbf{R}} - m_{kl}^{[i]}, \qquad \mathbf{G}_{kl}^{[i]} = \frac{\partial e_{kl}^{\tau_i\mathbf{R}}}{\partial \mathbf{R}}, \tag{151}$$

and $\Gamma^{[i]} = \sum_{\Upsilon} \xi_{kl}^{[i]} \mathbf{G}_{kl}^{[i]}$. The calculation of each $\mathbf{G}_{kl}^{[i]}$ would appear to require time proportional to N^3, thus leading to an overall complexity of $O(MN^3|\Upsilon|)$ (where $|\Upsilon|$ denotes the cardinality of Υ). We now show how this can be reduced to $O(M(LN^3 + |\Upsilon|))$ for some positive constant L.

Even though $\mathbf{R}$ and therefore $\exp(\tau_i\mathbf{R})$ and $\nabla\chi^2(\mathbf{R})$ are usually symmetric matrices in the application to NMR spectroscopy, we shall temporarily ignore symmetry in order to derive a compact formula for $\Gamma^{[i]}$ that holds for general matrices $\mathbf{A}$, and which may be interesting in its own right. Note that the restriction $k \le l$ is not needed in this case. First, by Eq. (145) we may write

$$\Gamma^{[i]} = \sum_{(k,l)\in\Upsilon} \xi_{kl}^{[i]} \mathbf{G}_{kl}^{[i]}(\mathbf{A}) = \sum_{(k,l)\in\Upsilon} \xi_{kl}^{[i]} \mathbf{D}_{\mathbf{E}_{kl}}\left(\tau_i, \mathbf{A}^{\mathrm{T}}\right). \tag{152}$$

If we now replace $\mathbf{D}_{\mathbf{E}_{kl}}(\tau_i, \mathbf{A}^{\mathrm{T}})$ by its integral representation, we get

$$\begin{aligned}\Gamma^{[i]} &= \sum_{(k,l)\in\Upsilon} \xi_{kl}^{[i]} \int_0^{\tau_i} e^{(\tau_i-\sigma)\mathbf{A}^{\mathrm{T}}} \mathbf{E}_{kl} e^{\sigma\mathbf{A}^{\mathrm{T}}}\, d\sigma \\ &= \int_0^{\tau_i} e^{(\tau_i-\sigma)\mathbf{A}^{\mathrm{T}}} \left(\sum_{(k,l)\in\Upsilon} \xi_{kl}^{[i]} \mathbf{E}_{kl} \right) e^{\sigma\mathbf{A}^{\mathrm{T}}}\, d\sigma = \int_0^{\tau_i} e^{(\tau_i-\sigma)\mathbf{A}^{\mathrm{T}}} \mathbf{V}^{[i]} e^{\sigma\mathbf{A}^{\mathrm{T}}}\, d\sigma,\end{aligned} \tag{153}$$

where the entries of the matrix $\mathbf{V}^{[i]}$ consist of the differences $\xi_{kl}^{[i]}$ between the measurement matrix and the matrix exponential at the mixing time τ_i for all pairs of indices $(k,l) \in \Upsilon$, and zero elsewhere. Regarding $\mathbf{V}^{[i]}$ as a direction matrix, one can write Eq. (153) as a *single* directional derivative in the direction $\mathbf{V}^{[i]}$, i.e.,

$$\Gamma^{[i]} = \mathbf{D}_{\mathbf{V}^{[i]}}\left(\tau_i, \mathbf{A}^{\mathrm{T}}\right), \tag{154}$$

so that $\nabla\chi^2(\mathbf{A})$ becomes a sum of M directional derivatives in M different directions:

$$\nabla\chi^2(\mathbf{A}) = \sum_{i=1}^{M} \Gamma^{[i]} = \sum_{i=1}^{M} \mathbf{D}_{\mathbf{V}^{[i]}}(\tau_i, \mathbf{A}^{\mathrm{T}}). \tag{155}$$

In the case of symmetric matrices such as the relaxation matrix $\mathbf{R}$, this formula is modified by Eq. (146) to obtain

$$\nabla\chi^2(\mathbf{R}) = \sum_{i=1}^{M} \mathrm{sym}(\mathbf{D}_{\mathbf{V}^{[i]}}(\tau_i, \mathbf{R})), \tag{156}$$

where $\mathbf{V}^{[i]}$ is a triangular matrix (because of the restriction $k \le l$).

Alternatively, we can reduce the amount of calculation and storage needed for this calculation by half by defining the symmetric matrix

$$\hat{\mathbf{V}}^{[i]} = \mathbf{V}^{[i]} + (\mathbf{V}^{[i]})^{\mathrm{T}} = \sum_{(k,l)\in\mathrm{T}} (1 + \delta_{kl})\xi_{kl}^{[i]}\hat{\mathbf{E}}_{kl}, \tag{157}$$

where $\hat{\mathbf{E}}_{kl}$ is an elementary symmetric matrix (see Section 6). Then, using Eq. (147), Eq. (155) is replaced by

$$\nabla\chi^2(\mathbf{R}) = \sum_{i=1}^{M} \left(\mathbf{D}_{\hat{\mathbf{V}}^{[i]}}(\tau_i, \mathbf{R}) - \frac{1}{2}\mathrm{diag}(\mathbf{D}_{\hat{\mathbf{V}}^{[i]}}(\tau_i, \mathbf{R})) \right). \tag{158}$$

7.2. *Summation Formulae for Multiple Mixing Times*

If $M > 1$ and one of the approximation/doubling methods described in Section 5 is used to compute $\chi^2(\mathbf{R})$ and $\nabla\chi^2(\mathbf{R})$, a great deal of time may be spent in two-sided iterative squaring. We now show that the time required to compute all the necessary exponentials and directional derivatives by this method can be reduced to some constant plus roughly $3M$ matrix multiplications, provided that the sequence of mixing times $\tau_1 < \tau_2 < \cdots < \tau_M$ forms an arithmetic (equidistant) or binary geometric progression. This should be compared to the time required by spectral methods once $\mathbf{R}$ has been diagonalized, which is roughly $5M$ matrix multiplications. Once again, it is easier to present these algorithms first for the general case of a nonsymmetric matrix $\mathbf{A}$.

First, suppose the mixing time sequence is an arithmetic progression

$$\tau_i = i\delta \qquad (i = 1, \ldots, M) \tag{159}$$

for some $\delta = \tau_1 > 0$. To simplify the presentation let $\mathbf{E} = \exp(\delta\mathbf{A})$, and let $\mathbf{D}^{[i]} = \mathbf{D}_{\mathbf{V}^{[i]}}(\delta, \mathbf{A}^{\mathrm{T}})$ be the directional derivatives evaluated at δ in each of

the directions $\mathbf{V}^{[i]}$. By means of the addition formula (53), the summation in Eq. (155) can be expanded into a triangular array as

$$\begin{aligned}
&\mathbf{D}^{[1]} \\
&+\mathbf{E}\mathbf{D}^{[2]} + \mathbf{D}^{[2]}\mathbf{E} \\
&+\mathbf{E}^2\mathbf{D}^{[3]} + \mathbf{E}\mathbf{D}^{[3]}\mathbf{E} + \mathbf{D}^{[3]}\mathbf{E}^2 \\
&\cdots \\
&+\mathbf{E}^{M-1}\mathbf{D}^{[M]} + \mathbf{E}^{M-2}\mathbf{D}^{[M]}\mathbf{E} + \cdots + \mathbf{E}\mathbf{D}^{[M]}\mathbf{E}^{M-2} + \mathbf{D}^{[M]}\mathbf{E}^{M-1},
\end{aligned} \tag{160}$$

where the sum in the ith row is equal to $\mathbf{D}_{\mathbf{V}^{[i]}}(\tau_i, \mathbf{A}^{\mathrm{T}})$. The following algorithm sums this array starting from the center of the bottom line:

Set: $\mathbf{C}_{\mathrm{L}} := \mathbf{0}$, $\mathbf{C}_{\mathrm{R}} := \mathbf{0}$, $\mathbf{S} := \mathbf{D}^{[M]}$ — if M is odd

Set: $\mathbf{C}_{\mathrm{L}} := \mathbf{E}\mathbf{D}^{[M]}$, $\mathbf{C}_{\mathrm{R}} := \mathbf{D}^{[M]}\mathbf{E}$, $\mathbf{S} := \mathbf{D}^{[M-1]} + \mathbf{C}_{\mathrm{L}} + \mathbf{C}_{\mathrm{R}}$ — if M is even

$\tilde{M} = \lfloor (M-1)/2 \rfloor$

Loop: $k = 1, \ldots, \tilde{M}$

$l = 2(\tilde{M} + 1 - k)$

$\mathbf{C}_{\mathrm{L}} := \mathbf{E}(\mathbf{D}^{[l]} + \mathbf{E}(\mathbf{D}^{[l+1]} + \mathbf{C}_{\mathrm{L}}))$

$\mathbf{C}_{\mathrm{R}} := (\mathbf{D}^{[l]} + (\mathbf{D}^{[l+1]} + \mathbf{C}_{\mathrm{R}})\mathbf{E})\mathbf{E}$

$\mathbf{S} := \mathbf{D}^{[l-1]} + \mathbf{C}_{\mathrm{L}} + \mathbf{C}_{\mathrm{R}} + \mathbf{E}\mathbf{S}\mathbf{E}$

End

At the end of this algorithm $\nabla\chi^2(\mathbf{A}) = \mathbf{S}$. Excluding the time needed to compute the matrices $\mathbf{D}^{[k]}$, this algorithm requires roughly $3M$ matrix multiplications.

If $\mathbf{A} = \mathbf{R}$ is a symmetric matrix, $\mathbf{E} = \exp(\delta\mathbf{R})$ and $\mathbf{D}^{[i]} = \mathbf{D}_{\hat{\mathbf{V}}^{[i]}}(\delta, \mathbf{R})$ is the directional derivative in the symmetric direction $\hat{\mathbf{V}}^{[i]}$, we replace the initialization of $\mathbf{S}$ above by $\mathbf{S} = \mathbf{D}^{[M]}/2$ if M is odd and by $\mathbf{S} = \mathbf{D}^{[M-1]}/2 + \mathbf{C}_{\mathrm{L}}$ if M is even. In addition, we omit calculation of $\mathbf{C}_{\mathrm{R}}$ and replace the calculation of $\mathbf{S}$ inside the loop by $\mathbf{S} = \mathbf{D}^{[l-1]}/2 + \mathbf{C}_{\mathrm{L}} + \mathbf{E}\mathbf{S}\mathbf{E}$. Then, at the end of the loop, the symmetric gradient is given by $\nabla\chi^2(\mathbf{R}) = \mathrm{sym}(\mathbf{S})$. This procedure takes about two-thirds the time required for the nonsymmetric case.

Suppose now that the mixing times form a geometric progression

$$\tau_i = 2^{i-1}\delta \qquad (i = 1, \ldots, M) \tag{161}$$

for some $\delta = \tau_1 > 0$. Then for general matrices **A**, Eq. (155) for $\nabla\chi^2(\mathbf{A})$ can be expanded by means of the doubling formula (54), as

$$\begin{gathered}
\mathbf{D}^{[1]} \\
\mathbf{E}\mathbf{D}^{[2]} + \mathbf{D}^{[2]}\mathbf{E} \\
\mathbf{E}(\mathbf{E}^2\mathbf{D}^{[3]} + \mathbf{D}^{[3]}\mathbf{E}^2) + (\mathbf{E}^2\mathbf{D}^{[3]} + \mathbf{D}^{[3]}\mathbf{E}^2)\mathbf{E} \\
\cdots \\
\ldots\left(\mathbf{E}^{2^{M-3}}(\mathbf{E}^{2^{M-2}}\mathbf{D}^{[M]} + \mathbf{D}^{[M]}\mathbf{E}^{2^{M-2}}) + (\mathbf{E}^{2^{M-2}}\mathbf{D}^{[M]} + \mathbf{D}^{[M]}\mathbf{E}^{2^{M-2}})\mathbf{E}^{2^{M-3}}\right)\ldots
\end{gathered} \tag{162}$$

This array can be summed by means of the following algorithm:

Set: $\mathbf{G} := \mathbf{D}^{[M]}$

$\mathbf{E}^{2^k} := (\mathbf{E}^{2^{k-1}})^2 \qquad (k = 1, \ldots, M-2)$

Loop: $k = 1, \ldots, M-1$

$\mathbf{G} := \mathbf{E}^{2^{M-1-k}}\mathbf{G} + \mathbf{G}\mathbf{E}^{2^{M-1-k}} + \mathbf{D}^{[M-k]}$

End

At the end of this algorithm $\nabla\chi^2(\mathbf{A}) = \mathbf{G}$. The computational complexity is $3M - 4$ matrix multiplications, not counting the time required to compute the directional derivatives. Unlike the algorithm for arithmetic progression, this algorithm requires the storage of the $M - 2$ binary powers of the matrix exponential **E**. The binary algorithm has the advantage of making it relatively easy to choose δ small enough that the initial directional derivatives can be evaluated without iterative squaring. In the case of a symmetric matrix such as **R**, all of the matrices involved in this algorithm are symmetric so that the time and storage can be reduced by half, and at the end of the loop $\nabla\chi^2(\mathbf{R}) = \mathbf{G} - \mathrm{diag}(\mathbf{G}/2)$.

APPENDIX: COMPUTATIONAL EXAMPLES

A1. The Spectral Product Representation

We begin by demonstrating the extent to which a partial spectral product decomposition as in Eq. (23) can improve the accuracy of the computed matrix exponential. Consider the following example of a stable matrix with a relatively wide eigenvalue separation but a poorly condi-

tioned eigenvector set:

$$\mathbf{A} = \mathbf{T}\mathbf{D}\mathbf{T}^{-1} = \begin{bmatrix} 1 & 10 & 100 \\ 1 & 9 & 100 \\ 1 & 11 & 99 \end{bmatrix} \begin{bmatrix} -0.001 & 0 & 0 \\ 0 & -1. & 0 \\ 0 & 0 & -100. \end{bmatrix} \begin{bmatrix} -209 & 110 & 100 \\ 1 & -1 & 0 \\ 2 & -1 & -1 \end{bmatrix}. \tag{163}$$

The condition numbers are $\kappa(\mathbf{A}) \approx 1.8793 \times 10^{10}$ and $\kappa(\mathbf{T}) \approx 4.4503 \times 10^4$, although $\kappa(\mathbf{T})$ can be reduced to $\kappa(\mathbf{T}) \approx 1.0637 \times 10^3$ by scaling the columns of $\mathbf{T}$.

When we compute $\exp(\mathbf{A})$ directly using the MATLAB function *expm* (based on diagonal Padé approximation for the exponential of order 6 and iterative squaring), and compare it to the exact result (with the norm $\|\exp(\mathbf{A})\|_2 \approx 435.885$), the norm of the absolute error is $\|(\exp(\mathbf{A})) - \exp(\mathbf{A})|_{\text{MATLAB}}\|_2 \approx 4.7135 \times 10^{-6}$. It should be noted that this loss in precision happens considerably before the occurrence of the maximal overshoot or hump (which occurs at about $t = 4$ with a value of $\|\exp(4\mathbf{A})\|_2 \approx 442.06$).

On the other hand, if we use a partial spectral decomposition to factor out $\exp(\lambda\mathbf{u}\mathbf{v}^T)$, where $\lambda = -100$ and $\mathbf{u}^T = [100\ 100\ 99]$, $\mathbf{v}^T = [2\ -1\ -1]$ are the corresponding left and right eigenvectors, then $\exp(\mathbf{A}) = (\mathbf{I} + (e^{-100} - 1)\mathbf{u}\mathbf{v}^T)e^{\mathbf{A} + 100\mathbf{u}\mathbf{v}^T}$. Since $\exp(-100) - 1 = -1$ in double precision, the first factor in the spectral product representation is computed to be

$$\mathbf{I} + (e^{-100} - 1)\mathbf{u}\mathbf{v}^T = \begin{bmatrix} -199 & 100 & 100 \\ -200 & 101 & 100 \\ -198 & 99 & 100 \end{bmatrix}, \tag{164}$$

while the exact spectral reduction of $\mathbf{A}$ is given by

$$\mathbf{A}[-100] = \mathbf{A} - (-100)\mathbf{u}\mathbf{v}^T = \begin{bmatrix} -9.791 & 9.89 & -0.1 \\ -8.791 & 8.89 & -0.1 \\ -10.791 & 10.89 & -0.1 \end{bmatrix}. \tag{165}$$

Note that the norm has decreased to $\|\mathbf{A}[-100]\|_2 \approx 24.188$ as compared to $\|\mathbf{A}\|_2 \approx 42306.222$. Computing $\exp(\mathbf{A}[-100])$ with the same function *expm* and multiplying it by the matrix in Eq. (164) now gives SP3 (partial spectral product representation based on the third eigenvalue), whose error norm $\|\exp(\mathbf{A}) - \exp(\mathbf{A})|_{\text{SP3}}\|_2 \approx 6.3038 \times 10^{-12}$ has been greatly reduced. If we compute $\exp(\mathbf{A})$ as the product of all three factors in the full spectral product (SP) representation, then the error norm is further reduced to $\|\exp(\mathbf{A}) - \exp(\mathbf{A})|_{\text{SP}}\|_2 \approx 2.5274 \times 10^{-13}$.

Finally, we calculate $e^{\mathbf{A}}|_J = \mathbf{T}e^{\mathbf{D}}\mathbf{T}^{-1}$ by means of the spectral sum representation, using the exact matrices in Eq. (163), and find that $\|e^{\mathbf{A}} - e^{\mathbf{A}}|_J\|_2 \approx 5.87 \times 10^{-14}$. If instead we use the MATLAB function *expm3* (which uses the eigenvalues and eigenvectors computed in double precision), we obtain $\|e^{\mathbf{A}} - e^{\mathbf{A}}|_{J_{\text{MATLAB}}}\|_2 \approx 1.5826 \times 10^{-6}$. Apparently the eigenvalues and eigenvectors have been corrupted by the poor condition of $\mathbf{T}$, even though the eigenvalues are not clustered.

When using partial spectral product decompositions as above, we have observed that even with Hermitian matrices somewhat higher precision should be used for the eigenvalues, eigenvectors, and their products $\lambda\mathbf{u}\mathbf{v}^{\mathrm{H}}$. Otherwise, the round-off error accumulation in computing $\mathbf{A}[\lambda_1, \lambda_2, \ldots, \lambda_k]$ may later be amplified in the product.

A2. Direct Padé versus the Approximation of $\mathbf{H}(\mathbf{B})$

The next example uses the same matrix $\mathbf{A}$ as that in Eq. (163) to demonstrate the extent to which the approximation of $\mathbf{H}(\mathbf{B})$ reduces the total number of matrix operations as compared to the direct Padé approximation of $\exp(\mathbf{B})$ (where $\mathbf{H}(\mathbf{B}) = \mathbf{B}\coth(\mathbf{B})$ and $\mathbf{B}$ denotes the scaled argument as in Section 2.2). The criterion for choosing a particular direct diagonal Padé approximation from this family is the same as that in Section 2.3, namely the minimization of the total number of matrix multiplications and divisions for the entire process. Each consecutive diagonal entry in the Padé table involves an additional matrix power, which means essentially an additional matrix multiplication. Furthermore, consecutive approximations are ε-accurate over increasing intervals of length, say $(\delta_1 < \delta_2 < \cdots < \delta_m)$, which are calculated from a sequence of nonlinear equations similar to Eq. (37). Therefore the first integer m at which the inequality $2\delta_m < \delta_{m+1}$ fails to be satisfied determines the optimal approximation. The reason is that from this point on iterative squaring (which takes one matrix multiplication to double the interval) becomes more efficient than what could be achieved by increasing the order of approximation, so that δ_m represents the optimal splitting of the interval between the local approximation process and iterative squaring.

For $\varepsilon = 2^{-52}$ we get $m = 5$, since $2\delta_5 > \delta_6$. Thus, the optimal double precision approximation is given by

$$\mathbf{E}_5(\mathbf{B}) = \frac{\mathbf{I} + \frac{1}{2}\mathbf{B} + \frac{1}{9}\mathbf{B}^2 + \frac{1}{72}\mathbf{B}^3 + \frac{1}{1008}\mathbf{B}^4 + \frac{1}{30240}\mathbf{B}^5}{\mathbf{I} - \frac{1}{2}\mathbf{B} + \frac{1}{9}\mathbf{B}^2 - \frac{1}{72}\mathbf{B}^3 + \frac{1}{1008}\mathbf{B}^4 - \frac{1}{30240}\mathbf{B}^5}, \tag{166}$$

which is valid over the interval $\delta = \delta_5 = 0.2981391$, and requires $5\frac{2}{3}$ matrix multiplications (given that one division equals $1\frac{2}{3}$ multiplications). With this value of δ the number of iterative squarings is $d = \lceil \ln_2(\|\mathbf{A}\|_2/\delta) \rceil =$

$\lceil \ln_2(42306.222/\delta) \rceil = 18$, so that the total number of matrix multiplications required by the direct Padé method is $23\frac{2}{3}$.

We now calculate the total number of matrix multiplications required by the method described in Section 2.3. The optimal approximation $\mathbf{H}_8(\mathbf{B})$ requires about $6\frac{2}{3}$ matrix multiplications and the optimal approximation interval is $\gamma = 1.1519224$. Hence the number of iterative squarings is $d = \lceil \frac{1}{2}\ln_2(\|\mathbf{A}^2\|/4\gamma) \rceil = 10$, for a total number of matrix multiplications of $16\frac{2}{3}$. The savings of seven matrix multiplications in this example amounts to about 30%.

The main reason for the increased efficiency in this example is that the error estimate for the approximation of $\mathbf{H}(\mathbf{B})$ is based on $\|\mathbf{A}^2\|$ rather than $\|\mathbf{A}\|$. If we consider only normal matrices for which $\|\mathbf{A}^2\| = \|\mathbf{A}\|^2$, the exponent d in the scale factor for our method can be written as $d = \lceil \log_2(\|\mathbf{A}\|/\xi) \rceil$, where $\xi = 2\sqrt{1.151922} = 2.1465533$ for double precision accuracy (see Section 2.3). The smaller optimal interval for the approximation (166) must be compensated for by additional $\log_2(\xi/\delta) = 2.84796$ iterative squarings. This means that even for normal matrices our method will save an average of 1.84796 matrix multiplications.

A3. Error Propagation in Two-Sided Iterative Squaring

We shall now illustrate the propagation of rounding error in the computation of the first directional derivative for the same matrix $\mathbf{A}$. Even for direction matrices of the same norm, the size of the derivative can vary greatly, as shown, for example, by $\|\mathbf{D}_{\mathbf{E}_{1,2}}(1,\mathbf{A})\|_2 \approx 85727.15885$ and $\|\mathbf{D}_{\mathbf{E}_{2,1}}(1,\mathbf{A})\|_2 \approx 43714.8932$. Both of these results were computed with the MAPLE program symbolic algebra using spectral representation formula (74) with 40 digits of accuracy.

We shall now demonstrate that in this example the propagation of rounding error in two-sided iterative squaring can be serious. Since $\|\mathbf{V}\|_2 = \|\mathbf{E}_{1,2}\|_2 = 1$, the optimal sinch approximation (see Section 5.1) for $\mathbf{D}_{\mathbf{V}}(1,\mathbf{A})$ is attained at $\nu = 3$ and $\beta = 0.05473601$. Therefore, the number of doubling steps is given by $d = \lceil \ln_2(\|\mathbf{A}\|_2/\beta) \rceil = 20$. We then calculate the optimal approximation of the directional derivative as well as the matrix exponential of the reduced matrix $\mathbf{A}/2^{20}$, using the MAPLE program with 40 digit precision, and *round both matrices* to a double precision accuracy of 17 significant digits. In order to evaluate the propagation of this initial rounding error, we carry out two-sided iterative squaring in MAPLE using 40 digit precision to avoid additional rounding error during the matrix multiplications, and denote the outcome by $\mathbf{D}_{\mathbf{V}}(1,\mathbf{A})|_{20}$. The norm of the difference from the exact (40 digit) result is $\|\mathbf{D}_{\mathbf{V}}(1,\mathbf{A})|_{20} - \mathbf{D}_{\mathbf{V}}(1,\mathbf{A})\|_2 \approx 2.17686 \times 10^{-4}$ with a relative accuracy of about 2.5×10^{-9}, indicating a substantial loss of significance due to the initial rounding error.

ACKNOWLEDGMENTS

We thank Professors Philip J. Davis and Lawrence Turyn for valuable comments concerning this paper. The first author (I.N.) also thanks Professor Nadav Liron. This work was supported by the NIH under Grants GM-38221 and GM-47467 and the NSF under Grant BIR-9115437 to the Harvard Medical School.

REFERENCES

1. M. Abramowitz and I. A. Stegun, "Handbook of Mathematical Functions," Dover, New York, 1972.
2. M. Athans and F. C. Schweppe, "Gradient Matrices and Matrix Calculations," MIT Lincoln Lab. Tech. Note 1965-53, Lexington, MA, 1965.
3. R. W. Atherton and A. E. De Gance, On evaluation of the derivative of the matrix exponential function, *IEEE Trans. Automat. Control* **20** (1975), 707–708.
4. H. F. Baker, On the integration of linear differential equations, *Proc. London Math. Soc.* **35** (1903), 333–374.
5. H. F. Baker, Alternants and continuous groups, *Proc. London Math. Soc. (2)* **3** (1904), 24–47.
6. R. Bellman, "Introduction to Matrix Analysis," 2nd ed., McGraw–Hill, New York, 1970.
7. B. A. Borgias, M. Gochin, D. J. Kerwood, and T. L. James, Relaxation matrix analysis of 2D NMR data, *Progr. NMR Spectrosc.* **22** (1990), 83–100.
8. J. W. Brewer, The gradient with respect to a symmetric matrix, *IEEE Trans. Automat. Control* **22** (1977), 265–267.
9. J. W. Brewer, The derivative of the exponential matrix with respect to a matrix, *IEEE Trans. Automat. Control* **22** (1977), 656–657.
10. J. W. Brewer, The derivative of the Riccati matrix with respect to a matrix, *IEEE Trans. Automat. Control* **22** (1977), 980–983.
11. J. W. Brewer, Kronecker products and matrix calculus in system theory, *IEEE Trans. Circuits Systems* **25** (1978), 772–781.
12. J. W. Brewer, Matrix calculus and the sensitivity analysis of linear dynamic systems, *IEEE Trans. Automat. Control* **23** (1978), 748–751.
13. Kuo-Tsai Chen, Integration of paths, geometric invariants and a generalized Baker–Hausdorff formula, *Ann. of Math.* **65**, No. 1 (1957), 163–178.
14. Kuo-Tsai Chen, Formal differential equations, *Ann. of Math.* **73**, No. 1 (1961), 111–133.
15. A. J. Chorin, T. J. R. Hughes, M. F. McCracken, and J. E. Marsden, Product formulas and numerical algorithms, *Comm. Pure Appl. Math.* **31** (1978), 205–256.
16. C. Davis, Explicit functional calculus, *Linear Algebra Appl.* **6** (1973), 193–199.
17. P. J. Davis, "Circulant Matrices," Wiley, New York, 1979.
18. A. S. Deif, Error bound for the equation $\mathbf{AX} + \mathbf{XB} = \mathbf{C}$, *Matrix and Tensor Quart.* (Dec. 1983/Mar. 1984), 66–73.
19. J. D. Dollard and C. N. Friedman, "Product Integration," Addison–Wesley, Reading, MA, 1979.
20. R. R. Ernst, G. Bodenhausen, and A. Wokaun, "Principles of Nuclear Magnetic Resonance in One and Two Dimensions," Oxford Univ. Press, Oxford, 1987.
21. A. E. Fekete, "A Real Linear Algebra," Dekker, New York, 1985.
22. R. P. Feynman, An operator calculus having applications in quantum electrodynamics, *Phys. Rev.* **84** (1951), 108–128.
23. P. Filipponi, An algorithm for computing functions of triangular matrices, *Computing* **26** (1981), 67–71.

24. H. Flanders, Analytic solutions of matrix equations, *Linear and Multilinear Algebra* **2** (1974), 241–243.
25. R. A. Frazer, W. J. Duncan, and A. R. Collar, "Elementary Matrices and Some Applications to Dynamics and Differential Equations," Macmillan Co., New York, 1946.
26. F. R. Gantmacher, "The Theory of Matrices," Vols. 1, 2, Chelsea, Bronx, NY, 1959.
27. A. Graham, "Kronecker Products and Matrix Calculus with Applications," Ellis Horwood, Chichester, 1981.
28. R. W. Hamming, *in* "Numerical Methods for Scientists and Engineers," 2nd ed., Chap. 39, McGraw–Hill, New York, 1973.
29. F. Hausdorff, Die symbolische Exponentialformel in der Gruppentheorie, *Ber. Sächsischen Akad. Wissenshaften (Math. Phys. Kl.)* **58** (1906), 19–48.
30. T. F. Havel, I. Najfeld, and J.-X. Yang, Matrix decompositions of two-dimensional nuclear magnetic resonance spectra, *Proc. Nat. Acad. Sci. U.S.A.* **91** (1994), 7962–7966.
31. P. Henrici, "Applied and Computational Complex Analysis," Vols. 1, 2, Wiley, New York, 1974, 1977.
32. R. A. Horn, The Hadamard product, *Proc. Sympos. Appl. Math.* **40** (1990), 87–169.
33. R. A. Horn and C. R. Johnson, "Topics in Matrix Analysis," Cambridge Univ. Press, Cambridge, UK, 1991.
34. R. I. Jennrich and P. B. Bright, Fitting systems of linear differential equations using computer generated exact derivatives, *Technometrics* **18** (1976), 335–392.
35. R. L. Jernigan, quoted in P. J. Flory and W. G. Miller, *J. Mol. Biol.* **15** (1966), 284–297.
36. R. Karplus and J. Schwinger, A note on saturation in microwave spectroscopy, *Phys. Rev.* **73** (1948), 1020–1026.
37. T. Kato, "Perturbation Theory for Linear Operators," 2nd ed., Springer-Verlag, Berlin/New York, 1976.
38. C. Kenney and A. J. Laub, Condition estimates for matrix functions, *SIAM J. Matrix Anal. Appl.* **10** (1989), 191–209.
39. P. Koehl and J.-F. Lefévre, The reconstruction of the relaxation matrix from an incomplete set of nuclear Overhauser effects, *J. Magnetic Resonance* **86** (1990), 565–583.
40. C. Lanczos, *in* "Applied Analysis," Chap. IV, Sect. 23, Prentice Hall, Englewood Cliffs, NJ, 1956.
41. M. Lutzky, Parameter differentiation of exponential operators and the Baker–Campbell–Hausdorff formula, *J. Math. Phys.* **9** (1968), 1125–1128.
42. W. Magnus, On the exponential solution of differential equations for a linear operator, *Commun. Pure Appl. Math.* **7** (1954), 649–673.
43. M. Marcus and H. Minc, "A Survey of Matrix Theory and Matrix Inequalities," Dover, New York, 1992.
44. R. Mathias, Evaluating the Fréchet derivative of the matrix exponential, *Numer. Math.* **63** (1992), 213–226.
45. C. B. Moler and C. F. Van Loan, Nineteen dubious ways to compute the exponential of a matrix, *SIAM Rev.* **20** (1978), 801–836.
46. E. Nuding, Schrankentreue Berechnung der Exponentialfunktion wesentlich-nichtnegativer Matrizen, *Computing* **26** (1981), 57–66.
47. G. Opitz, Steigungsmatrizen, *Z. Angew. Math. Mech.* **44** (1964), T52–T54.
48. B. N. Parlett, A recurrence among the elements of functions of triangular matrices, *Linear Algebra Appl.* **14** (1976), 117–121.
49. B. N. Parlett, and K. C. Ng, Accurate computation of the matrix exponential, *in* "Computing Methods in Applied Sciences and Engineering, Vol. VII" (R. Glowinsky and J. L. Lions, Eds.), pp. 107–116, Amsterdam/New York, 1986.
50. M. C. Pease III, "Methods of Matrix Algebra," Academic Press, New York, 1965.

51. J. B. Plant, On the computation of transient matrices for time invariant systems, *Proc. IEEE* **56** (1968), 1397–1398.
52. D. L. Powers, On the differentials of certain matrix functions, *Canad. J. Math.* **23**, No. 2 (1971), 282–286.
53. D. L. Powers, The Fréchet differential of a primary matrix function, *Canad. J. Math.* **25**, No. 3 (1973), 554–559.
54. R. F. Reinhart, Extension of the derivative concept for functions of matrices, *Proc. Amer. Math. Soc.* **8** (1957), 329–335.
55. R. F. Reinhart, Differential of a primary matrix function, *Rend. Circ. Math. Palermo* **15** (1966), 209–215.
56. J. R. Roche, On the sensitivity of the matrix exponential problem, *RAIRO Anal. Numér.* **15**, No. 3 (1981), 249–255.
57. P. C. Rosenbloom, Bounds on functions of matrices, *Amer. Math. Monthly* **74** (1967), 920–926.
58. R. F. Snider, Perturbation variation methods for a quantum Boltzmann equation, *J. Math. Phys.* **5** (1964), 1580–1587.
59. E. U. Stickel, An algorithm for fast high precision computation of matrix exponential and logarithm, *Analysis* **10** (1990), 85–95.
60. E. U. Stickel, On the Fréchet derivative of a matrix function, *Linear Algebra Appl.* **91** (1987), 83–88.
61. C. F. Van Loan, The sensitivity of the matrix exponential, *SIAM J. Numer. Anal.* **14** (1977), 971–981.
62. C. F. Van Loan, Computing integrals involving the matrix exponential, *IEEE Trans. Automat. Control* **23** (1978), 395–404.
63. W. J. Vetter, Derivative operations on matrices, *IEEE Trans. Automat. Control* **15** (1970), 241–244.
64. W. J. Vetter, An extension to gradient matrices, *IEEE Trans. Systems Man Cybernet.* **1** (1971), 184–186.
65. W. J. Vetter, Matrix calculus operations and Taylor expansions, *SIAM Rev.* **15** (1973), 352–369.
66. G. Wagner, S. G. Hyberts, and T. F. Havel, NMR structure determination in solution: A critique and comparison with X-ray crystallography, *Ann. Rev. Biophys. Biomol. Struct.* **21** (1992), 167–198.
67. G. Walz, Computing the matrix exponential and other matrix functions, *J. Comput. Appl. Math.* **21** (1988), 119–123.
68. R. C. Ward, Numerical computation of the matrix exponential with accuracy estimates, *SIAM J. Numer. Anal.* **14** (1977), 600–610.
69. E. Wegert and L. N. Trefethen, From the Buffon needle to the Kreiss matrix theorem, *Amer. Math. Monthly* **101** (1994), 133–139.
70. R. M. Wilcox, Exponential operators and parameter differentiation in quantum physics, *J. Math. Phys.* **8** (1967), 962–982.
71. K. Wüthrich, "NMR of Proteins and Nucleic Acids," Wiley, New York, 1986.
72. P. F. Yip and D. A. Case, A new method for refinement of macromolecular structures based on nuclear Overhauser effect spectra, *J. Magnetic Resonance* **83** (1989), 643–648.
73. P. F. Yip, A computationally efficient method for evaluating the gradient of 2D NOESY intensities, *J. Biomol. NMR* **3** (1993), 361–365.