High-Order Time-Stepping for Nonlinear PDE through Rapid Estimation of Block Gaussian Quadrature Nodes

James V. Lambers

Department of Mathematics
University of Southern Mississippi
Hattiesburg, MS 39401 USA

Abstract. The stiffness of systems of ODEs that arise from spatial discretization of PDEs causes difficulties for both explicit and implicit time-stepping methods. Krylov Subspace Spectral (KSS) methods present a balance between the efficiency of explicit methods and the stability of implicit methods by computing each Fourier coefficient from an individualized approximation of the solution operator of the PDE. While KSS methods are explicit methods that exhibit a high order of accuracy and stability similar to that of implicit methods, their efficiency needs to be improved. A previous asymptotic study of block Lanczos iteration yielded estimates of extremal block Gaussian quadrature nodes for each Fourier component and led to an improvement in efficiency. Here, a more detailed asymptotic study is performed in order to rapidly estimate *all* nodes, thus drastically reducing computational expense without sacrificing accuracy.

Exponential propagation iterative (EPI) methods provide an efficient approach to the solution of large stiff nonlinear systems of ODE, compared to standard integrators. However, the bulk of the computational effort in these methods is due to products of matrix functions and vectors, which can become very costly at high resolution due to an increase in the number of Krylov projection steps needed to maintain accuracy. Here, it is proposed to modify EPI methods by using KSS methods, instead of standard Krylov projection methods, to compute products of matrix functions and vectors. Numerical experiments demonstrate that this modification causes the number of Krylov projection steps to become bounded independently of the grid size, thus dramtically improving efficiency and scalability.

Keywords: spectral methods, Gaussian quadrature, variable-coefficient, block Lanczos method, stability, heat equation
PACS: 02.60.Lj, 02.60.Dc, 02.60.Jh

INTRODUCTION

Consider an autonomous, stiff system of ODE

$$\mathbf{y}' = F(\mathbf{y}), \quad \mathbf{y}(t_0) = \mathbf{y}_0, \tag{1}$$

such as one that would arise from spatial discretization of a PDE. Stiffness poses problems for both explicit and implicit time-stepping methods. For explicit methods, the time step is severely restricted, while for implicit methods, an ill-conditioned system must be solved during each time step, for which an iterative method requires many iterations or a specially developed preconditioner [14].

Exponential propagation iterative (EPI) methods, due to Tokman, et al. [12, 14] are designed to reduce the number of Krylov projection steps needed to compute any required matrix function-vector products of the form $\mathbf{w} = f(A\tau)\mathbf{b}$, where f is a function, A is an ill-conditioned matrix, τ is a parameter determined by the time step, and $\mathbf{b}$ is a vector. One approach to computing $\mathbf{w}$ for a given symmetric matrix A is to apply the Lanczos algorithm to A with the initial vector $\mathbf{b}$, until we obtain a matrix X_j with orthonormal columns and a tridiagonal matrix T_j such that $X_j^T A X_j = T_j$. Then, we can compute the approximation

$$\mathbf{w}_j = \|\mathbf{b}\|_2 X_j f(T_j\tau)\mathbf{e}_1, \tag{2}$$

where $\mathbf{e}_1 = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}^T$. Since each column $\mathbf{x}_k$ of the matrix X_j is of the form $\mathbf{x}_k = p_{k-1}(A)\mathbf{b}$, where $p_n(A)$ is a polynomial of degree n in A, then $\mathbf{w}_j$ is the product of a polynomial in A of degree $j-1$ and $\mathbf{b}$. Since the matrix A arises from a stiff PDE, the eigenvalues of A are not clustered, which means that a large number of Lanczos iterations might be required in order to obtain a good approximation of $\mathbf{w}$.

The difficulty that time-stepping methods have with stiffness is due to the coupling of the components of the solution with widely varying frequencies. Another problem is the use of the same function, whether polynomial or rational, to approximate all of these components of $f(A\tau)\mathbf{b}$, when such a function cannot effectively approximate $f(\lambda\tau)$ on a large

interval except at high degree, which is computationally expensive. A solution to this problem is to use Krylov subspace spectral (KSS) methods [9, 10], which is a component-wise approach in which each Fourier coefficient of the solution is computed using an approximation of the solution operator that is tailored to that component. Specifically, KSS methods use an interpolating polynomial with frequency-dependent interpolation points to approximate the function f. As a result, these methods demonstrate a high order of accuracy and stability like that of implicit methods.

Untill now, KSS methods have been used mainly in linear PDE on n-dimensional boxes, for $n = 1,2,3$, with either periodic or homogeneous boundary conditions. A succesful implementation of KSS methods for nonlinear PDEs was used by Guidotti, et al. when a one-node KSS method was applied to nonlinear diffusion equations from image processing [5] to obtain first-order accuracy in time. However, in order to achieve higher-order accuracy for nonlinear PDE, in addition to using more nodes, it is also necessary to account for the nonlinearity more carefully than with a simple linearization at each time step. This can be accomplished by combining KSS methods with EPI methods.

KRYLOV SUBSPACE SPECTRAL METHODS

To review the essential aspects of KSS methods [9], we consider the parabolic PDE $u_t + Lu = 0$ on the interval $[0, 2\pi]$, where L is a Sturm-Liouville operator, with appropriate initial conditions and periodic boundary conditions. The idea behind KSS methods is that the Fourier coefficients of the solution $\tilde{u}(x, t_{n+1})$ are obtained by applying the exact solution operator to the previously computed solution $\tilde{u}(x, t_n)$. These Fourier coefficients are given by

$$\hat{u}(\omega, t_{n+1}) = \left\langle \frac{1}{\sqrt{2\pi}} e^{i\omega x}, e^{-L\Delta t}\tilde{u}(x, t_n) \right\rangle, \tag{3}$$

where $\langle \cdot, \cdot \rangle$ denotes the standard inner product on $[0, 2\pi]$ and $e^{-L\Delta t}$ is the solution operator of the PDE.

As a result of the spatial discretization of (3), we obtain the following bilinear form

$$\mathbf{u}^T f(A)\mathbf{v}, \tag{4}$$

where $\mathbf{u} = \frac{1}{\sqrt{2\pi}} e^{i\omega x}$ and $\mathbf{v} = \tilde{u}(x, t_n)$ are N-vectors, $A = L_N$, where L_N is a spectral discretization of L, and $f(\lambda) = e^{-\lambda t}$.

Since the matrix A is symmetric positive definite, it has real eigenvalues $b = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N = a > 0$, and corresponding orthonormal eigenvectors $\mathbf{q}_j$, $j = 1, \ldots, N$. As a result, we have the following spectral decomposition

$$\mathbf{u}^T f(A)\mathbf{v} = \sum_{j=1}^{N} f(\lambda_j)\mathbf{u}^T \mathbf{q}_j \mathbf{q}_j^T \mathbf{v}. \tag{5}$$

As mentioned by Golub and Meurant in [3], (3) can also be viewed as a Rieman-Stieltjes integral

$$\mathbf{u}^T f(A)\mathbf{v} = \int_a^b f(\lambda)\, d\alpha(\lambda), \tag{6}$$

where the measure $d\alpha(\lambda)$ is derived from the coefficients of $\mathbf{u}$ and $\mathbf{v}$ in the basis of eigenvectors. This allows approximation using Gaussian quadrature rules, where the nodes and weights are obtained using the Lanczos algorithm applied to A with initial vectors $\mathbf{u}$ and $\mathbf{v}$ [3].

In the case where $\mathbf{u} \neq \mathbf{v}$,the presence of a negative weight would destabilize the quadrature rule [1]. Alternatively, we consider the approximation of the 2×2 matrix integral

$$\begin{bmatrix} \mathbf{u} & \mathbf{v} \end{bmatrix}^T f(A) \begin{bmatrix} \mathbf{u} & \mathbf{v} \end{bmatrix}. \tag{7}$$

We use the most general K-node quadrature formula, as described in [3], to get an approximation for (6) of the form

$$\int_a^b f(\lambda)\, d\mu(\lambda) = \sum_{j=1}^{2K} f(\lambda_j)\mathbf{v}_j \mathbf{v}_j^T + error, \tag{8}$$

where, for each j, λ_j is a scalar and $\mathbf{v}_j$ is a 2-vector. Each node λ_j is an eigenvalue of the matrix

$$\mathscr{T}_K = \begin{bmatrix} M_1 & B_1^T & & \\ B_1 & M_2 & B_2^T & \\ & \ddots & \ddots & \ddots \\ & & B_{K-1} & M_K \end{bmatrix}, \tag{9}$$

which is a block-tridiagonal matrix of order $2K$. The vector $\mathbf{v}_j$ consists of the first two elements of the corresponding normalized eigenvector. The matrices M_j and B_j are computed using the block Lanczos algorithm [4]:

The block KSS method starts by defining

$$R_0(\omega) = \left[\begin{array}{cc} \hat{\mathbf{e}}_\omega & \mathbf{u}^n \end{array} \right],$$

where $\hat{\mathbf{e}}_\omega$ is a discretization of $\frac{1}{\sqrt{2\pi}} e^{i\omega x}$ and $\mathbf{u}^n$ is the computed solution at time t_n. The QR factorization of $R_0(\omega)$ yields $R_0(\omega) = X_1(\omega)B_0(\omega)$. Then, block Lanczos iteration is applied to the discretized operator L_N with initial block $X_1(\omega)$, producing a block tridiagonal matrix $\mathscr{T}_K(\omega)$ of the form (9), where each entry is a function of ω. Then, each Fourier coefficient of the solution at time t_{n+1} can be expressed as

$$[\hat{\mathbf{u}}^{n+1}]_\omega = \left[B_0^H E_{12}^H e^{-\mathscr{T}_K(\omega)\Delta t} E_{12} B_0 \right]_{12}, \quad E_{12} = \left[\begin{array}{cc} \mathbf{e}_1 & \mathbf{e}_2 \end{array} \right]. \tag{10}$$

This algorithm has temporal accuracy $O(\Delta t^{2K-1})$ for parabolic problems [9]. Even higher-order accuracy, $O(\Delta t^{4K-2})$, is obtained for the second-order wave equation [10]. Furthermore, under appropriate assumptions on the coefficients of the PDE, the 1-node KSS method is unconditionally stable [9, 10].

ASYMPTOTIC ANALYSIS OF BLOCK LANCZOS ITERATION

Let $\mathbf{u}$ be a discretization of the solution u on a uniform N-point grid. Then, KSS methods use the initial block $R_0 = \left[\begin{array}{cc} \hat{\mathbf{e}}_\omega & \mathbf{u} \end{array} \right]$, for each $\omega = -N/2+1, \ldots, N/2$. We then use block Lanczos iteration to compute the block tridiagonal matrix $\mathscr{T}_K$ in (9). In [13] it is shown that every (nonzero) off-diagonal entry of M_j or B_j, for $j = 1, 2, \ldots$, is a Fourier coefficient of some function that is a differential operator applied to u. Therefore, as long as the Fourier coefficients of u decay to zero at a sufficiently high rate as $|\omega| \to \infty$, these off-diagonal entries will also decay to zero.

It follows that in this high-frequency limit, the block tridiagonal matrix $\mathscr{T}_K$ produced by block Lanczos applied to R_0 as defined above converges to the matrix that would be obtained by applying "non-block" Lanczos iteration to the two columns of R_0 *separately*, and then alternating rows and columns of the tridiagonal matrices produced by these iterations. Therefore, by reordering the rows and columns of $\mathscr{T}_K$ in such a way that odd-numbered and even-numbered rows and columns are grouped together, we find that the eigenvalue problem for this matrix *decouples*, and the block Gaussian quadrature nodes can be obtained by computing the eigenvalues of these smaller, tridiagonal matrices [13]. For finite ω, we can then use non-block Lanczos to at least estimate the true block Gaussian quadrature nodes. In [13], it is shown that this analysis also applies in the case where the leading coefficient p is *not* constant.

This decoupling reveals that we can obtain approximations of half of the block Gaussian quadrature nodes for all Fourier components by applying "non-block" Lanczos iteration to the matrix L_N with initial vector $\mathbf{u}^n$, the computed solution, as is done in standard Krylov projection methods such as those described in [6, 7, 8]. These nodes will be referred to as frequency-independent nodes. To estimate the other half of the nodes, we perform an asymptotic analysis of Lanczos iteration applied to L_N with initial vector $\hat{\mathbf{e}}_\omega$; these are called frequency-dependent nodes.

In [13] we considered the case where the matrix A comes from a spectral discretization of the operator $Lu = -pu_{xx} + q(x)u$, where p is a constant, and assumed periodic boundary conditions. Carrying out three iterations, which corresponds to a fifth-order accurate KSS method for a parabolic PDE, yields the following recursion coefficients as functions of the wave number ω, after neglecting lower-order terms:

$$\left[\begin{array}{ccc} \alpha_1 & \overline{\beta_1} & 0 \\ \beta_1 & \alpha_2 & \overline{\beta_2} \\ 0 & \beta_2 & \alpha_3 \end{array} \right] \approx \left[\begin{array}{ccc} p\omega^2 & \|\tilde{\mathbf{q}}\|_2 & 0 \\ \|\tilde{\mathbf{q}}\|_2 & p\omega^2 & 2p|\omega|\|\mathbf{q}_x\|_2/\|\tilde{\mathbf{q}}\|_2 \\ 0 & 2p|\omega|\|\mathbf{q}_x\|_2/\|\tilde{\mathbf{q}}\|_2 & p\omega^2 \end{array} \right].$$

It follows that the nodes can easily be estimated as

$$\lambda_{1,\omega} = p\omega^2, \quad \lambda_{i,\omega} = p\omega^2 \pm \sqrt{\beta_1^2 + \beta_2^2}, \quad i = 2, 3. \tag{11}$$

KSS-EPI METHODS

The combination of KSS and EPI methods is easily described: whenever an EPI method computes a matrix function-vector product of the form $\varphi(A\tau)\mathbf{b}$, for some function φ, matrix A, scaling parameter τ based on the time step, and vector $\mathbf{b}$, the following procedure is carried out in place of standard Krylov projection as in (2):

1. Use an FFT to decompose $\mathbf{b} = \mathbf{b}_L + \mathbf{b}_H$, where $\mathbf{b}_L$ consists of low-frequency components and $\mathbf{b}_H$ contains all other (high-frequency) components.
2. Use standard Krylov projection as in (2) to compute $\varphi(A\tau)\mathbf{b}_L$.
3. Use KSS, with nodes prescribed as in the previous section, to compute $\varphi(A\tau)\mathbf{b}_H$.
4. Add the results of steps 2 and 3 to obtain $\varphi(A\tau)\mathbf{b}$.

The decomposition of $\mathbf{b}$ in step 1 must be chosen so that step 2 does not require many more Krylov projection steps than would be required for step 3, which is one more than the desired order of temporal accuracy. However, it is also important to not include too few low-frequency components in $\mathbf{b}_L$, as the nodes prescribed in the previous section are based on a high-frequency analysis and are therefore not effective choices at low frequencies [13]. Step 3 can be implemented in such a way that $\lceil \frac{q+1}{2} \rceil$ FFTs are performed, where q is the desired order of temporal accuracy [2].

SUMMARY

When solving stiff systems of nonlinear ODE derived from PDE, an increase in the number of grid points in the spatial discretization of the PDE does not necessarily require a corresponding increase in the number of Krylov projection steps needed to maintain high-order accuracy in time. By employing a componentwise approach, in which each component of the solution with respect to an appropriate orthonormal basis is computed using an individualized approximation of the function f, the Krylov subspace dimension can be bounded independently of the grid size and instead determined by the desired temporal order of accuracy. The effective approach is demonstrated through numerical experiments in [2].

REFERENCES

1. Atkinson, K.: *An Introduction to Numerical Analysis, 2nd Ed.* Wiley (1989)
2. Cibotarica, A., Lambers, J. V., Palchak, E.: High-Order Time-Stepping for Nonlinear PDE through Rapid Estimation of Block Gaussian Quadrature Nodes. Submitted.
3. Golub, G. H., Meurant, G.: Matrices, Moments and Quadrature. *Proceedings of the 15th Dundee Conference*, June-July 1993, Griffiths, D. F., Watson, G. A. (eds.), Longman Scientific & Technical (1994)
4. Golub, G. H., Underwood, R.: The block Lanczos method for computing eigenvalues. *Mathematical Software III*, J. Rice Ed., (1977) 361-377.
5. Guidotti, P., Kim, Y., Lambers, J. V.: Image Restoration with a New Class of Forward-Backward Diffusion Equations of Perona-Malik Type. *SIAM Journal on Imaging Sciences* **6**(3) (2013) 1416-1444.
6. Hochbruck, M., Lubich, C.: A Gautschi-type method for oscillatory second-order differential equations, *Numer. Math.* **83** (1999) 403-426.
7. Hochbruck, M., Lubich, C.: On Krylov Subspace Approximations to the Matrix Exponential Operator. *SIAM J. Numer. Anal.* **34** (1996) 1911-1925.
8. Hochbruck, M., Lubich, C., Selhofer, H.: Exponential Integrators for Large Systems of Differential Equations. *SIAM J. Sci. Comput.* **19** (1998) 1552-1574.
9. Lambers, J. V.: Enhancement of Krylov Subspace Spectral Methods by Block Lanczos Iteration. *Electron. T. Numer. Ana.* **31** (2008) 86-109.
10. Lambers, J. V.: An Explicit, Stable, High-Order Spectral Method for the Wave Equation Based on Block Gaussian Quadrature. *IAENG Journal of Applied Mathematics* **38** (2008) 333-348.
11. Lambers, J. V.: Implicitly Defined High-Order Operator Splittings for Parabolic and Hyperbolic Variable-Coefficient PDE Using Modified Moments. *International Journal of Computational Science* **2** (2008) 376-401.
12. Loffeld, J., Tokman, M.: Comparative Performance of Exponential, Implicit and Explicit Integrators for Stiff Systems of ODEs. Submitted.
13. Palchak, E. M., Cibotarica, A., Lambers, J. V.: Solution of Time-Dependent PDE Through Rapid Estimation of Block Gaussian Quadrature Nodes. Submitted.
14. Tokman, M.: Efficient integration of large stiff systems of ODEs with exponential propagation iterative (EPI) methods. *J. Comp. Phys.* **213** (2006) 748-776.