Commun. Comput. Phys.
doi: 10.4208/cicp.OA-2021-0182

Vol. **32**, No. 1, pp. 156-188
July 2022

Reliability Investigation of BiCGStab and IDR Solvers for the Advection-Diffusion-Reaction Equation

Chris Schoutrop[1,*], Jan ten Thije Boonkkamp[2] and Jan van Dijk[1]

[1] *Department of Applied Physics, Eindhoven University of Technology, The Netherlands.*
[2] *Department of Mathematics and Computer Science, Eindhoven University of Technology, The Netherlands.*

Received 8 September 2021; Accepted (in revised version) 14 April 2022

Abstract. The reliability of BiCGStab and IDR solvers for the exponential scheme discretization of the advection-diffusion-reaction equation is investigated. The resulting discretization matrices have real eigenvalues. We consider BiCGStab, IDR(S), BiCGStab(L) and various modifications of BiCGStab, where S denotes the dimension of the shadow space and L the degree of the polynomial used in the polynomial part. Several implementations of BiCGStab exist which are equivalent in exact arithmetic, however, not in finite precision arithmetic. The modifications of BiCGStab we consider are; choosing a random shadow vector, a reliable updating scheme, and storing the best intermediate solution. It is shown that the Local Minimal Residual algorithm, a method similar to the ''minimize residual" step of BiCGStab, can be interpreted in terms of a time-dependent advection-diffusion-reaction equation with homogeneous Dirichlet boundary conditions for the residual, which plays a key role in the convergence analysis. Due to the real eigenvalues, the benefit of BiCGStab(L) compared to BiCGStab is shown to be modest in numerical experiments. Non-sparse (e.g. uniform random) shadow residual turns out to be essential for the reliability of BiCGStab. The reliable updating scheme ensures the required tolerance is truly achieved. Keeping the best intermediate solution has no significant effect. Recommendation is to modify BiCGStab with a random shadow residual and the reliable updating scheme, especially in the regime of large Péclet and small Damköhler numbers. An alternative option is IDR(S), which outperforms BiCGStab for problems with strong advection in terms of the number of matrix-vector products. The MATLAB code used in the numerical experiments is available on GitLab: `https://gitlab.com/ChrisSchoutrop/krylov-adr`, a C++ implementation of IDR(S) is available in the Eigen linear algebra library: `http://eigen.tuxfamily.org`.

AMS subject classifications: 00A79, 70-08, 15A06, 15B05, 15A18

*Corresponding author. *Email addresses:* `c.e.m.schoutrop@tue.nl` (C. Schoutrop), `j.h.m.tenthijeboonkkamp@tue.nl` (J. ten Thije Boonkkamp), `j.v.dijk@tue.nl` (J. van Dijk)

Key words: BiCGStab, IDR, shadow residual, advection-diffusion-reaction equation.

1 Introduction

Advection-diffusion-reaction (ADR) equations and their discrete approximations are ubiquitous in the modeling of physical systems [1–5]. A wide variety of discretization schemes for the ADR equation, such as finite difference, (pseudo)spectral, finite element and finite volume methods exist. For the solution to be representative, the discretization scheme must not only be convergent in the limit of infinitesimally fine grids, but also yield representative results for more pragmatic grid sizes. A counterexample, is the discretization of the ADR equation in the presence of strong advection where the central differencing scheme yields spurious oscillations if the grid is too coarse [6, p. 83]. This discrepancy is reflected by the eigenvalues of the exact ADR-operator and the discretized operator, i.e., the exact operator has *real* eigenvalues, whereas the discretized version has *complex* eigenvalues. However, for the exponential discretization scheme of [6, 7] which we use here, the discretized version also has real eigenvalues.

After the discretization step a linear system is obtained which must be solved to obtain the approximate solution. Such linear systems are of the type $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$, with $\boldsymbol{A}$ a generally sparse, asymmetric, but invertible matrix of size $N\times N$. In this paper we mainly consider 3D equations. Note that even with a modest $M=102$ grid points per direction this results in a linear system with $N=10^6$ unknowns. A robust method for solving such linear systems is by factorizing $\boldsymbol{A}$ into a pair of lower and upper triangular matrices using the well-known LU decomposition [8, p. 96], and subsequently computing $\boldsymbol{x}$ by solving two triangular systems using backward and forward substitution. The main downside of LU decomposition is that for a sparse system matrix there can be significant fill-in; the factors $\boldsymbol{L}$ and $\boldsymbol{U}$ are not guaranteed to be sparse. As a result the time complexity of factorizing the resulting $\boldsymbol{A}$ for discretized three-dimensional ADR-equations is in general $\mathcal{O}(N^{7/3})$ [9].

Another approach are iterative methods, most commonly the Krylov subspace methods such as the Conjugate Gradient (CG) method. Such methods seek successive approximations to the solution as a projection on the linear subspace $\mathcal{K}_k(\boldsymbol{A},\boldsymbol{r}_0)=\text{span}\{\boldsymbol{r}_0,\boldsymbol{A}\boldsymbol{r}_0,\boldsymbol{A}^2\boldsymbol{r}_0,\cdots,\boldsymbol{A}^{k-1}\boldsymbol{r}_0\}$ with $\boldsymbol{r}_0:=\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}_0$ the residual generated by some initial guess $\boldsymbol{x}_0$. For linear systems obtained from discretizing a second order PDE with a grid spacing proportional to $N^{-1/3}$, the CG method requires $\mathcal{O}(N^{4/3})$ flops to reach a given tolerance ϵ [9], saving a factor N compared to the LU decomposition. A second benefit is that the memory requirement of $\mathcal{O}(N)$ is modest, as only the matrix itself and a constant number of vectors of size N have to be stored. Additionally, CG is an optimal Krylov method in the sense that in each iteration the error is minimized over the $\boldsymbol{A}$-norm [10].

The main drawback of CG is the requirement of a symmetric, positive definite matrix $\boldsymbol{A}$. For more general invertible matrices the Faber-Manteuffel theorem [11] states