BOOK OF ABSTRACTS

2019 INTERNATIONAL CONFERENCE ON PRECONDITIONING TECHNIQUES FOR SCIENTIFIC AND INDUSTRIAL APPLICATIONS

JULY 1-3, 2019
UNIVERSITY OF MINNESOTA, TWIN CITIES
ABSTRACTS: INVITED TALKS
Linear preconditioning is an established research area, as it is testified by the 20th anniversary of this conference. Non-linear preconditioning is more recent and much less understood. A standard approach to obtain a good linear preconditioner is to start with a good stationary iterative method, like domain decomposition or multigrid, and then to use the associated operator as a preconditioner for a Krylov method, or in other words to use the Krylov method as an accelerator. Since the Krylov method finds in general a better polynomial than the stationary iterative method alone, it is worthwhile using Krylov acceleration in the linear case. In the non-linear case, one can proceed similarly: to obtain a good non-linear preconditioner, one starts with an effective fixed point iteration for solving the non-linear problem, like a non-linear domain decomposition method or the full approximation scheme in multigrid, and then instead of just running the fixed point iteration, one uses Newton’s method to solve the preconditioned system at the fixed point. My first goal is to explain in detail how to obtain such non-linear preconditioners for Newton’s method. My second goal is to show that while there are many cases where using Newton’s method is beneficial to accelerate convergence, there are also cases where using Newton’s method is detrimental, in great contrast to the linear case and Krylov methods. Newton’s method can destroy convergence, and I will show a domain decomposition example where using Newton’s method to accelerate the non-linear domain decomposition method leads to period doubling and chaos, while the domain decomposition method alone converges very well.

\[^1\text{University of Geneva}\]
When approximating PDEs with the finite element method, large sparse linear systems must be solved. The ideal preconditioner yields convergence that is algorithmically optimal and parameter robust, i.e. the number of Krylov iterations required to solve the linear system to a given accuracy does not grow substantially as the mesh or problem parameters are changed.

Achieving this for the stationary Navier-Stokes has proven challenging: LU factorisation is Reynolds-robust but scales poorly with degree of freedom count, while Schur complement approximations such as PCD and LSC degrade as the Reynolds number is increased.

Building on the ideas of Schöberl, Benzi & Olshanskii, and Lee et al. [1, 2, 3], in this talk we present the first preconditioner for the Newton linearisation of the stationary Navier–Stokes equations in three dimensions that achieves both optimal complexity and Reynolds-robustness. The scheme combines augmented Lagrangian stabilisation to control the Schur complement, a divergence-capturing additive Schwarz relaxation method on each level, and a specialised prolongation operator involving non-overlapping local Stokes solves. The properties of the preconditioner are tailored to the divergence-free Scott–Vogelius discretisation [4].

We present 3D simulations with over one billion degrees of freedom with robust performance from Reynolds numbers 10 to 5000.

References


[1] Mathematical Institute, University of Oxford
[2] Department of Computer Science, Durham University
[3] Department of Computer Science, University of Chicago
The numerical simulation of poroelastic problems [1] has become of increasing importance due to their wide range of applications. Intensive research has been focused on the design of efficient methods for solving the large linear systems arising from the discretization of Biot’s model, since in real simulations it is the most consuming part. There are two typical approaches for solving such problem: fully-coupled or monolithic methods and iterative coupling methods. In the context of monolithic techniques, we aim to design efficient and robust preconditioners [2, 3] to accelerate the convergence of Krylov subspace methods. The proposed block preconditioners for solving the Biot’s model are based on the well-posedness of the obtained discrete systems [4, 5], and are robust with respect to both physical and discretization parameters. This latter is a very important feature of the proposed solvers since in many physical applications the values of the parameters involved in the model vary over orders of magnitude. Numerical results are presented to support the theoretical results.

References


The numerical solution of high frequency wave propagation has been a longstanding challenge in computational science and engineering. This talk discusses some recent developments in designing efficient preconditioners for time-harmonic wave equations. We consider a sequence of examples with important applications, and for each we construct an efficient preconditioner (approximate inverse) that allows one to solve the system with a small number of iterations. From these examples emerges a new framework, where sparsity, geometry of wave phenomenon, and highly accurate discretizations are combined together to address this challenging topic.
Large-scale linear least-squares (LS) problems occur widely in practice. They arise both in their own right and as subproblems of more general nonlinear problems. Moreover, the normal equations that are naturally connected to LS problems can shed light on solving the Schur complement systems that are routinely faced in many applications.

Our focus in this talk is on gaining a better understanding of the case when the sparse LS problem contains additional coupling terms represented by one or more dense rows. It has long been recognised that the effectiveness of sparse matrix techniques for directly solving such problems is severely limited by the presence of dense rows. We consider, in particular, the following $m \times n$ ($m > n$) LS problem

$$\min_x \|Ax - b\|_2 = \min_x \left\| \begin{pmatrix} A_s & A_d \end{pmatrix} x - \begin{pmatrix} b_s \\ b_d \end{pmatrix} \right\|_2,$$

in which each row of the $m_d \times n$ block $A_d$ is considered to be dense and $A_s$ is $m_s \times n$ with $m_s \gg m_d \geq 1$; the vector $b$ is partitioned conformally. These LS problems represent a simple motivating case for more general situations that appear in practice where a dense substructure hidden in the problem may prohibit efficient solution.

There are a number of ways to tackle this problem. Classical approaches based on direct methods are summarized in the monograph [2]; see also [4]. More recently, methods based on preconditioned iterative methods [5] or Schur complement reduction [6] have been considered. In this talk, we discuss a number of approaches. One specific approach discussed here is based on matrix stretching in which dense rows are replaced by submatrices with much sparser rows [3] [1]. Experimental problems demonstrate not only the strengths of stretching but also some of its limitations [8]; these point towards future research directions.

References


---

1Faculty of Mathematics and Physics, Charles University, Prague
2The University of Reading and STFC Rutherford Appleton Laboratory, UK


A factorization based framework for building scalable algebraic preconditioners

X.S. Li

Factorization based preconditioning algorithms, most notably incomplete LU (ILU) factorization, have been shown to be robust and applicable to wide ranges of problems. However, traditional ILU algorithms are not amenable to scalable implementation. In recent years, we have seen a lot of investigations using low-rank compression techniques to build approximate factorizations. A key to achieving lower complexity is the use of hierarchical matrix algebra, stemming from the $\mathcal{H}$- and $\mathcal{H}^2$-matrix research. In addition, the multilevel algorithm paradigm provides a good vehicle for a scalable implementation. The goal of this talk is to give an overview of the various hierarchical matrix formats, such as hierarchically semi-separable matrix (HSS), hierarchically off-diagonal low-rank matrix (HODLR) and butterfly matrix, and to explain the algorithm differences and approximation quality. We will illustrate many practical issues of these algorithms using our parallel libraries STRUMPACK and ButterflyPACK, and demonstrate their effectiveness and scalability while solving the very challenging problems, such as high frequency wave equations.

\footnote{Lawrence Berkeley National Laboratory}
Many important engineering and scientific systems require the solution of time-dependent PDE systems. Many of these systems have specific stability needs in order to compute realistic solutions such as needing A-stable or L-stable methods. For example, the Eddy Currents Equation is a stiff parabolic PDE with a possible rapid transient, and benefits from an L-stable method. Similarly, the incompressible Navier–Stokes equations for fluid flow are differential algebraic equations (DAE), and can be viewed as infinitely stiff. Therefore, L-stable time stepping methods can be beneficial there as well.

Certain classes of implicit Runge-Kutta (IRK) time-stepping methods, such as the Radau I and Radau II methods provide L-stability, but one price of using an IRK method is needing to solve large linear systems at each time step. Suppose, for example, our PDE has been linearized and discretized with \( N \) degrees of freedom. Using an \( s \)-stage IRK method leads to an \( sN \times sN \) linear system that must be solved at each time step. These systems are block-(\( s \times s \)) systems, where each block is \( N \times N \).

In this talk, we investigate preconditioners for such systems. In some cases, we can take advantage of known structure in the subblocks. For example, in the time-dependent incompressible Navier–Stokes equations, each subblock is related to a linear system from the steady-state fluid flow equations, for which there are several effective preconditioners.

---

1Texas Tech University, USA (victoria.howle@ttu.edu)
2Texas Tech University, USA (katharine.long@ttu.edu, Md-Masud.Rana@ttu.edu)
3Adams State University, USA (ameek@adams.edu)
Preconditioned iterative methods are usually the only feasible solution approach for very large scale computations. As such, the use of advanced hardware is often an important issue. Next-generation computers are generally based on a significant improvement of peak performances at a relatively low energy cost through an increasing level of concurrency of the computations. The efficient porting of preconditioners for discretized PDEs on such machines, however, is not simply a matter of implementation tricks, but often requires to re-think the underlying algorithm, or to develop new ones, in order to take full advantage of the larger computing parallel degree. This scenario opens up new perspectives on the development of next-generation preconditioners, which should look for scalability and inherent parallelism as fundamental paradigms for their formulation. This Mini-Symposium aims to be forward-looking, presenting ideas and methodology for preconditioning that may be suitable for the next generation of computers.

Speakers and tentative titles:

1. Daniel Osei-Kuffuor (CASC, Livermore Nat Lab, USA), "Algebraic multigrid reduction for coupled systems of PDEs"

2. Carlo Janna (University of Padova, Italy), "An adaptive AMG preconditioner for modern High Performance Computers"

3. Edmond Chow (Georgia Tech, USA), "Iterative methods in biochemistry"

4. Andy Wathen (Oxford University, UK), "Parallel preconditioning for time-dependent PDEs"
Monolithic (or all-at-once) discretizations of evolutionary problems most often give rise to non-symmetric linear(ised) systems of equations which can be of very large dimension for PDE problems. In this talk we will describe preconditioners for such systems with guaranteed fast convergence via use of MINRES (not LSQR or CGNE) or GMRES. These results apply with standard time-stepping schemes and relate to Toeplitz matrix technology and preconditioning via circulants using the FFT. Simple parallel computational results will be shown for the heat equation and the wave equation.

This is joint work with Elle McDonald (CSIRO, Australia), Jennifer Pestana (Strathclyde University, UK) and Anthony Goddard (Durham University, UK)
In a wide variety of probabilistic simulations, it is necessary to compute sample vectors from a multivariate Gaussian distribution with zero mean and a given covariance matrix. For large-scale problems, this task is computationally expensive, and despite the availability of several approaches, fast methods are still highly desired since such sampling remains a computational bottleneck in these simulations.

For a symmetric positive definite covariance matrix $A$, the canonical method of constructing a Gaussian sample $y \sim N(0, A)$ is to compute $y = Sz$, where $A = SS^T$ and $z$ is a standard normal vector. With this definition of $S$, it is clear that the covariance of $y$ will be the matrix $A$. As a result, any $S$ that satisfies $A = SS^T$ can be used, for example, the lower triangular Cholesky factor or principal square root of $A$. Use of the Cholesky factor is most common.

Instead of Cholesky factorization, we focus on a category of sampling methods based on matrix polynomials. Here, sample vectors of the form $p(A)z$ are computed, where $p$ is a polynomial chosen such that $p(A)$ approximates the principal square root of $A$. One need not form the matrix $p(A)$ explicitly. Instead, for each $z$, $p(A)z$ is computed from a sequence of products with the matrix $A$. Two main types of polynomial methods have been used: those that choose $p(A)$ as an expansion of Chebyshev or other orthogonal polynomials, and those that construct the sample from a Krylov subspace.

In this talk, we show how these polynomial sampling methods can be preconditioned. Unlike the linear system case, preconditioning here changes the problem being solved. However, a sample with the desired covariance can still be recovered.

Consider a factorization $G^T G$ that approximates $A^{-1}$, which we may call a preconditioner. Now consider using the preconditioned matrix $GAG^T$, rather than $A$, in a polynomial sampling method, to produce a sample

$$\tilde{w} = p(GAG^T)z$$

which is approximately distributed as $N(0, GAG^T)$. Since $GAG^T$ is well-conditioned, we expect that only a small number of terms is required in the polynomial approximation to the square root of $GAG^T$. To construct a sample with the desired covariance, apply $G^{-1}$ to each sample $\tilde{w}$,

$$\tilde{y} = G^{-1}\tilde{w} = G^{-1}p(GAG^T)z$$

which is approximately distributed as $N(0, A)$. Such a sample approximates $G^{-1}(GAG^T)^{1/2}z$. The matrix $S = G^{-1}(GAG^T)^{1/2}$ satisfies

$$SS^T = G^{-1}(GAG^T)^{1/2}(GAG^T)^{1/2}G^{-T} = A$$

---

1 Georgia Institute of Technology
2 University of Minnesota
as desired, but it is not a Cholesky factor or square root of $A$. The idea leading to the method described here is that $S$ can be any of an infinite number of quantities that satisfies $SS^T = A$. How well the covariance of $\tilde{y}$ approximates $A$ depends on the accuracy of $p$ and not on $G$. The convergence rate depends on the quality of the approximation $G^T G \approx A^{-1}$. For example, in the extreme case when we have an exact Cholesky factorization at our disposal, then $G^T G = A^{-1}$, and we only need to take $p(A) = I$, i.e., $\tilde{y}$ becomes $\tilde{y} = G^{-1} p(GAG^T) z = G^{-1} z$. This corresponds to the standard method based on Cholesky ($G^{-1}$ is the lower triangular Cholesky factor of $A$). However, we now have the option of using an approximate factorization instead of an exact one.

As usual, the preconditioned matrix $GAG^T$ is not formed explicitly. Instead, $G$ and $G^T$ are applied to vectors in these methods. To construct the desired sample, however, we must be able to easily solve with $G$. (The roles of matrix-vector multiplications and solves are reversed if we define the preconditioner $GG^T$ as an approximation to $A$.) These requirements are more demanding than the usual requirements for preconditioners.
(MS) An adaptive AMG preconditioner for modern High Performance Computers

C. Janna

A. Franceschini

V. A. Paludetto Magri

The numerical simulation of modern engineering problems can easily incorporate millions or even billions of unknowns. In several applications, sparse linear systems with symmetric positive definite matrices need to be solved, and algebraic multigrid (AMG) methods represent common choices for the role of iterative solvers or preconditioners. The reason for their popularity relies on the fast convergence that these methods provide even in the solution of large size problems, which is a consequence of the AMG ability to reduce particular error components across their multilevel hierarchy. Despite carrying the name ”algebraic”, most of these methods still make assumptions on additional information other than the global assembled matrix, such as the knowledge of the operator’s near kernel, which limits their applicability as black-box solvers. In this talk, we introduce a novel AMG approach [1] featuring the adaptive factored sparse approximate inverse (aFSAI) method as a flexible smoother and, following the ideas of adaptive AMG methods [2, 3], uses Krylov-based eigensolvers to uncover the system matrix near-kernel. A novel greedy strategy is finally introduced to adaptively compute a sparse prolongation operator able to approximate optimal prolongation. It will be also shown that all of the algorithms needed to form such an approach are suitable for modern many core architecture, such as GPUs [4], which are expected to be the building blocks for the exascale computing. The performance of the proposed AMG is verified through the solution of a set of model problems along with real-world engineering test cases [5].

References


1University of Padova

2Stanford University

3University of Padova

Simulation of reservoir geomechanics and multiphase flow involves solving multi-physics problems in which multiphase flow is tightly coupled with geomechanical processes. To capture this dynamic interplay, fully implicit methods, also known as monolithic approaches, are usually preferred. The main bottleneck of this strategy is the cost of solving the linear systems resulting from discretization of the problem. Because of the strong coupling present in the continuous problem, efficient techniques such as algebraic multigrid (AMG) cannot be directly applied to the discrete linear systems. In this work, we present our efforts in developing an algebraic framework based on multigrid reduction that is suited for tightly coupled systems of PDEs. Using this framework, the decoupling between the equations is done algebraically through defining appropriate interpolation and restriction operators. One can then employ existing solvers for each of the decoupled blocks or design a new solver based on knowledge of the physics. We will demonstrate the applicability of our framework to multiphase flow coupled with geomechanics. We show that the framework is flexible to accommodate a wide range of scenarios, as well as efficient and scalable for large problems.
Preconditioning is a field that studies the acceleration of iterative solutions of a linear (or non-linear) system in large-scale scientific applications. Machine learning is a discipline that uses computer algorithms to perform tasks, relying on patterns, data, and inference rather than explicit instructions. The emerging intersection of the two gives rise to new perspectives and opportunities for solving important problems in the respective areas. This minisymposium is comprised of three talks that develop preconditioning techniques for computing machine learning models such as Gaussian processes, canonical correlation analysis, and linear discriminant analysis; and a talk on using machine learning techniques for choosing the most effective preconditioner from a portfolio of general purpose preconditioners. These problems exemplify the synergy of preconditioning and machine learning and motivate interests in the community for expanding the frontiers.

The four talks of this minisymposium are:

1. Lucas Erlandson (Georgia Tech). “Preconditioning for Fast Solves in Gaussian Processes Accelerated by Hierarchical Matrices”.


3. Boris Shustin (Tel Aviv University). “Randomized Riemannian Preconditioning for Quadratically Constrained Problems”.

4. Jie Chen (IBM Research). “Preconditioner Selection for General Linear Systems by Using Neural Networks”.

IBM Research
Iterative methods for solving linear systems suffer the ill conditioning of the matrix; hence, preconditioning is crucial to the success of an iterative solution. Over the years, a number of general-purpose preconditioners were developed, but none works equally well for all problems. Selecting the right preconditioner (possibly with parameters) is more often an art. We take a machine learning approach and compute a model that predicts the solution time of a linear system for a certain combination of iterative method and preconditioner. Specifically, the model is parameterized by using a neural network, commonly regarded as a universal functional approximator in the limit. We present preliminary results and discuss pros and cons of the approach, hinting on a promising direction for theorists and practitioners.
In this talk will discuss the problem of spectrum approximation of large matrices in the context of matrix function trace estimation. We discuss two different approaches to approximate the matrix spectrum, namely the Lanczos quadrature approach and a histogram approach which uses Preconditioned Stochastic Variance Reduction Gradient (SVRG) method. The topics of this talk are detailed in [1, 2].

Understanding the spectrum of a matrix $A \in \mathbb{R}^{n \times n}$ is a fundamental task in countless applications. In matrix-matrix multiplication time, it is possible to the complete spectrum of the matrix. However, little is known about algorithms that break this runtime barrier. One such context where the matrix spectrum information is needed is the problem of estimating the trace of matrix functions. This problem appears in applications ranging from machine learning, scientific computing, to computational biology. In this talk, we first present an inexpensive method to estimate the trace of $f(A)$ for cases where $f$ is analytic inside a closed interval and $A$ is a symmetric positive definite matrix. The method combines three key ingredients, namely, the stochastic trace estimator, Gaussian quadrature, and the Lanczos algorithm. As examples, we consider the problems of estimating the log-determinant ($f(t) = \log(t)$), the Schatten $p$-norms ($f(t) = t^{p/2}$), the Estrada index ($f(t) = e^t$) and the trace of matrix inverse ($f(t) = t^{-1}$). The multiplicative and additive error bounds for the approximations obtained by this method will be presented. In addition, we present error bounds for other useful tools such as approximating the log-likelihood function in the context of maximum likelihood estimation of Gaussian processes. We also discuss numerical experiment results illustrating the performance of the method on different problems arising from various applications.

The above Lanczos quadrature approach will be inexpensive for large sparse matrices, however, its theoretical runtime will depend on the condition number of the matrix. In the second part of the talk, we will discuss a histogram approach to approximate the matrix spectrum whose runtime is independent of the matrix condition number, and faster than matrix multiplication time. Using tools from stochastic trace estimation, polynomial approximation, and fast system solvers, we show how to efficiently isolate different ranges of $A$'s spectrum and approximate the number of singular values in these ranges. We thus effectively compute a histogram of the spectrum, which can stand in for the true singular values in many applications. We use this primitive to give the first algorithms for approximating a wide class of symmetric matrix norms in faster than matrix multiplication time. For example, we give a $(1 + \epsilon)$ approximation algorithm for the Schatten-1 norm (the nuclear norm) running in just $\tilde{O}(nnz(A) n^{1/3} + n^2 \epsilon^{-3})$ time for $A$ with uniform row sparsity or $\tilde{O}(n^{2.18} \epsilon^{-3})$ time for dense matrices. The runtime scales smoothly for general Schatten-$p$ norms, notably becoming $\tilde{O}(p.nnz(A) \epsilon^{-3})$ for any $p \geq 2$. At the same time, we show that the complexity of spectrum approximation is inherently tied to fast matrix multiplication in the small $\epsilon$ regime. We prove that achieving milder $\epsilon$ dependencies in
our algorithms would imply faster than matrix multiplication time triangle detection for general graphs. This further implies that highly accurate algorithms running in subcubic time yield subcubic time matrix multiplication. As an application of our bounds, we show that precisely computing all effective resistances in a graph in less than matrix multiplication time is likely difficult, barring a major algorithmic breakthrough.

References


Gaussian Process based models provide much inferential power, as well as variance information, however they require the solution of dense linear systems with size proportional to the number of points, making the direct approach infeasible for large problems. This talk discusses the use of preconditioners for hierarchical matrix based solutions to accelerate Gaussian Process prediction and optimization. The use of the SMASH $\mathcal{H}^2$ hierarchical matrix from [2] results in near linear matrix-vector products, which can be used in iterative solvers. We consider and present results for a variety preconditioners including both geometry aware and global preconditioners.

Given a symmetric positive definite kernel $K : K(\cdot, \cdot) \in \mathbb{R}$ with training points $X$, test points $X_*$, and training and test output vectors $f$ and $f_*$ respectively, we have the multivariate Gaussian distribution of the output

$$
\begin{bmatrix}
f \\
f_*
\end{bmatrix} \sim \mathcal{N}
\left(0,
\begin{bmatrix}
K(X,X) & K(X,X_*) \\
K(X_*,X) & K(X_*,X_*)
\end{bmatrix}
\right)
$$

This results in the predictive distribution

$$
f_*|X, f, X_* \sim \mathcal{N}(\text{mean, covariance})
$$

$$
\text{mean} : K(X_*, X)(K(X, X))^{-1} y
$$

$$
\text{covariance} : K(X_*, X_*) - K(X_*, X)(K(X, X))^{-1} K(X, X_*)
$$

In Gaussian Process literature, the direct way of performing a Gaussian Process prediction involves computing the solutions of the linear system via the Cholesky decomposition of the $K(X, X)$ covariance matrix, an $O(n^3)$ operation [3]. This $O(n^3)$ complexity makes Gaussian Process regression prohibitively expensive for large problems, and as such there is great motivation to develop methods which decrease the complexity. The most popular kernel is the Gaussian Radial Basis, or Squared Exponential, kernel

$$
K(x, y) = e^{-\frac{1}{2} \frac{|x-y|^2}{\sigma^2}},
$$
which is smooth, and thus the covariance matrix is a good candidate for representation via a hierarchical matrix.

The underlying idea of hierarchical matrices is that, for data sparse matrices, the matrix can be approximated via a hierarchy of matrices with a low-rank representation, and using properties of this approximation operations can be greatly accelerated. We use the SMASH $\mathcal{H}^2$ matrix which provides parallelizable, near-linear matrix-vector products which can be used in matrix-free methods. While using hierarchical matrices for Gaussian Process Regression [1] and Kernel Ridge Regression [4] have been investigated previously, this talk focuses on the use of preconditioners in conjunction with iterative solvers accelerated by hierarchical matrices.

The need for preconditioners is due to the ill conditioning of the covariance matrix. One of the steps of the SMASH matrix construction is the construction of a cluster tree, which provides information on the geometry of the problem, which we able to exploit in our geometry aware preconditioners. We investigate the use of LU, SVD, and SVD with dynamic stabilization inside of this geometry aware framework. In addition to these preconditioners we also investigate and present results for Chebyshev polynomial preconditioners of varying degrees.

References


1 Introduction

In this talk, we consider optimization problems with quadratic equality constraints. That is, we consider problems of the form

$$\min f(x_1, \ldots, x_k) \text{ s.t. } x_i^T B_i x_i = 1 \ (i = 1, \ldots, k)$$

where $B_1, \ldots, B_k$ are fixed symmetric positive definite (SPD) matrices. Such optimization problem with quadratic equality constraints are prevalent in machine learning. Indeed, two important examples are CCA and LDA:

1. Suppose $A$ and $B$ are two SPD matrices. The eigenvector corresponding to the largest generalized eigenvalue is the solution of $\max x^T A x \text{ s.t. } x^T B x = 1$. One example application is LDA, in which $B$ is the Gram matrix of a shifted version of the input data.

2. Let $X \in \mathbb{R}^{n \times d_x}$ and $Y \in \mathbb{R}^{n \times d_y}$ be two data matrices, and let $\lambda_x, \lambda_y \geq 0$ be two regularization parameters. In CCA we seek to find a $u \in \mathbb{R}^{d_x}$ and a $v \in \mathbb{R}^{d_y}$ that maximize $u^T X^T Y v$ subject to $u^T (X^T X + \lambda_x I_{d_x}) u = 1$ and $v^T (Y^T Y + \lambda_y I_{d_y}) v = 1$.

Unfortunately, methods for solving the aforementioned problems typically involve computing matrix inverses and decomposition. In the context of machine learning, these matrices are actually Gram matrices of input data matrices, and as such the computations are too expensive for large scale datasets.

In the talk, which is based on a recent paper [3], we propose a sketching based approach for solving Canonical Correlation Analysis (CCA) and Linear Discriminant Analysis (LDA) that reduces the cost dependence on the input size. The proposed algorithms feature randomized preconditioning combined with Riemannian optimization.

2 Riemannian Optimization

The main idea in [3] is to use Riemannian optimization where the preconditioner is incorporated as the metric. We developed some Riemannian components for the ellipsoid geometry necessary
for the application of Riemannian optimization in our problem settings based on [1]. These components naturally extended to the product manifold setting (see also Section 3.1.6 in [1]) for the CCA problem. We present the analytical expressions:

- **Notation:** given $B \in \mathbb{R}^{d \times d}$, a SPD matrix, the $d - 1$ dimensional ellipsoid is defined by
  \[ S^B = \left\{ x \in \mathbb{R}^d : x^T B x = 1 \right\}. \]

- **Retraction:** the retraction mapping allows us to take step on the tangent space at a point $x \in S^B$:
  \[ R_x(x_i) = \frac{x + x_i}{\|x + x_i\|_B}. \]  
  (5)

- **Vector Transport:** is useful for manipulating tangent vectors from two different tangent spaces:
  \[ T_{\eta_x}(x_i) x_i = \frac{1}{\|x + \eta_x\|_B} \left[ I - \frac{(x + \eta_x)(x + \eta_x)^T B}{\|x + \eta_x\|_B^2} \right] x_i. \]  
  (6)

- **Riemannian Metric:** is endowed by the inner product in the ambient space $\mathbb{R}^d$: for a constant matrix $M$ we have
  \[ g_x(x_i, \eta_x) = x_i^T M \eta_x. \]  
  That makes it a Riemannian submanifold of $\mathbb{R}^d$. The metric also serves as the preconditioner.

- **Orthogonal Projection on the Tangent Space:** given a Riemannian metric we have
  \[ P_x = (I_n - (x^T B M^{-1} B x)^{-1} M^{-1} B x x^T B). \]  
  (7)

We also developed analytical expressions for the Riemannian gradient and the Riemannian Hessian in ambient coordinates. For example for the CCA problem:

- **The inputs are** two data matrices $X \in \mathbb{R}^{n \times d_x}$ and $Y \in \mathbb{R}^{n \times d_y}$, and two regularization parameters $\lambda_x, \lambda_y \geq 0$. We denote $d = d_x + d_y$, and $z = \begin{bmatrix} u \\ v \end{bmatrix} \in \mathbb{R}^d$ where $u \in \mathbb{R}^{d_x}$ and $v \in \mathbb{R}^{d_y}$.

- **The constraint set:** is a product manifold of two ellipsoids $S_{xy} = S_{\Sigma_{xx}} \times S_{\Sigma_{yy}}$.

- **The objective function:** to be minimized is
  \[ f(z) = -\frac{1}{2} z^T \begin{bmatrix} 0 & \Sigma_{xy} \\ \Sigma_{xy}^T & 0 \end{bmatrix} z. \]  
  (8)

- **The Riemannian gradient and the Riemannian Hessian in ambient coordinates:**
  \[ \text{grad} f(z) = P_z M^{-1} \nabla f(z) = -\begin{bmatrix} P_u M_{xx}^{-1} \Sigma_{xy} v \\ P_v M_{yy}^{-1} \Sigma_{xy}^T u \end{bmatrix} \]  
  \[ \text{Hess} f(z)[\eta_z] = P_z M^{-1} \begin{bmatrix} (u^T M_{xx} P_u M_{xx}^{-1} \Sigma_{xy} v) \cdot \Sigma_{xx} & -\Sigma_{xy} \\ -\Sigma_{xy}^T & (v^T M_{yy} P_v M_{yy}^{-1} \Sigma_{xy}^T u) \cdot \Sigma_{yy} \end{bmatrix} \eta_z \]  
  (10)
3 Results

Our main theoretical result is a bound on the condition number of the Riemannian Hessian at the optimum: $\frac{2\sigma_1}{\sigma_1 - \sigma_2} \cdot \kappa(\Sigma, M)$ (for CCA), which exemplifies how the preconditioner affects the (asymptotic) convergence. The costs of evaluating the Riemannian components depend on the number of non-zeros of the data matrices and the costs of computing operations with the metric matrix and its inverse. If we use CountSketch transformation to form the preconditioner, specific bound and costs are calculated. More details, results, and numerical experiments for the CCA and LDA problem can be found in the full paper [3].

Finally, we present some numerical experiments for the CCA and LDA with the MEDIAMILL and COVTYPE datasets implemented using MANOPT [2]. The preconditioning strategy is implemented with the CountSketch transform and the optimization methods are Riemannian Conjugate Gradient (CG) and Riemannian Trust Region.

Figure 1: Results for CCA on MEDIAMILL (left) and LDA on COVTYPE (right).

Full details, including additional experimental results, see [3].

References


3 Other transformations such as Subsampled Randomized Hadamard Transform may also be a good choice
4 Datasets were downloaded for libsvm’s website: https://www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/
Preconditioners for Saddle Point Problems is an active research area. An example of such a problem is the linearized and discretized incompressible Navier Stokes equations. One has to solve block-structured indefinite linear systems. The successful design of robust, scalable, and efficient preconditioners is intimately connected with an understanding of the structure of the resulting block matrix system. Effective preconditioners are often based on an approximate block decomposition of the discretized incompressible Navier Stokes equations. This requires a careful consideration of the spectral properties of the component block operators and their Schur complement operators. Through this purely algebraic view of preconditioning, a simplified system of block component equations is developed. Inclusion of "physics based" preconditioners of the various parts can lead to effective preconditioners with optimal or nearly optimal convergence rates for academic and industrial problems.

1. *The LDLT factorization and new extensions for saddle point system preconditioning.* by Jos Maubach and Wil Schilders

2. *Efficient and robust preconditioners for high Reynolds number laminar flows.* by Xin He and Kees Vuik

3. *Application of block preconditioners to isogeometric analysis discretizations of the incompressible Navier-Stokes equations.* by Hana Hornikova, Kees Vuik, and Jiri Egermaier

4. *Flexible computational abstractions for complex preconditioners* by Lawrence Mitchell, Patrick E. Farrell, Rob C. Kirby, Matthew G. Knepley, and Florian Wechsung

---

1 TU Eindhoven
2 TU Delft
The augmented Lagrangian (AL) preconditioner [1], belonging to the class of block structured preconditioners [6, 7], is originally proposed to solve saddle point systems arising from the incompressible Navier-Stokes equations discretized by the finite element method (FEM). The AL preconditioner features a purely algebraic construction and robustness with respect to the Reynolds number and mesh refinement. Because of these attractive features, recent research was devoted to the further development and extension of the AL preconditioner, notably the extension [3] to the context of stabilized finite volume methods (FVM), which are widely used in industrial computational fluid dynamic applications, and the modified variant [2] with reduced computational complexities.

In [4] we introduce an alternative method to approximate the Schur complement for the AL preconditioner, which leads to a new variant of the AL preconditioner. This new method approximates the Schur complement through its inverse form and facilitates the utilization of the existing Schur complement approximations. Among the available candidates, the Schur complement approximation from the SIMPLE preconditioner [5] is chosen and substituted into the inverse Schur complement approximation for the AL preconditioner. The so-arising new variant of the AL preconditioner reduces the number of Krylov subspace iterations by a factor up to 36 compared to the original one on the turbulent applications of the maritime industry.

Since the new method to approximate the Schur complement for the AL preconditioner use the existing Schur complement approximations, the following questions straightforwardly raise. Does the utilization of other existing Schur complement approximations deliver a better performance than that from the SIMPLE preconditioner? If so, which Schur complement approximation is the most efficient one? Does the optimal choice depend on the test problem and parameters arising from the physics and discretization, e.g. the Reynolds number and grid size? To answer these questions, in this talk we utilize the existing Schur complement approximations not only from the SIMPLE preconditioner but also from the LSC and PCD operators to construct the new Schur complement approximation in the AL preconditioner. Extensive comparisons between the considered Schur complement approximations are carried out on a wide range of numerical experiments in the laminar context to evaluate the effect of the Reynolds number, mesh anisotropy and refinement on the optimal choice. Furthermore, the advantage over the traditional Schur complement approximation in terms of the reduced number of the Krylov subspace iterations is exhibited. This work is expected to provide an alternative to efficiently apply the AL preconditioner for solving saddle point systems arising

---

1Institute of Computing Technology, Chinese Academy of Sciences, No.6 Kexueyuan South Road Zhongguancun Haidian District Beijing, P.R. China 100190.
2Delft Institute of Applied Mathematics, Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands.
from the Navier-Stokes equations and provide a fundamental guideline for the more complicated
turbulent flow calculations.

References


ers for the incompressible Navier-Stokes equations. International Journal for Numerical


[4] X. He, C. Vuik and C.M. Klaij. Combining the augmented Lagrangian preconditioner with
the SIMPLE Schur complement approximation. SIAM Journal on Scientific Computing,

[5] C.M. Klaij and C. Vuik. SIMPLE-type preconditioners for cell-centered, colocated finite
volume discretization of incompressible Reynolds-averaged Navier-Stokes equations. Inter-


Small block overlapping, and non-overlapping, Schwarz methods are theoretically highly attractive as multilevel smoothers for a wide variety of problems that are not amenable to point relaxation methods. Examples include monolithic Vanka smoothers for Stokes, overlapping vertex-patch decompositions for $H(\text{div})$ and $H(\text{curl})$ problems [1], along with nearly incompressible elasticity [4], and augmented Lagrangian schemes [2, 3].

While it is possible to manually program these different schemes, their use in general purpose libraries has been held back by a lack of generic, composable interfaces. We present a new approach to the specification and development such additive Schwarz methods in PETSc that cleanly separates the topological space decomposition from the discretisation and assembly of the equations. Our preconditioner is flexible enough to support overlapping and non-overlapping additive Schwarz methods, and can be used to formulate line, and plane smoothers, Vanka iterations, amongst others. I will illustrate these new features with some examples utilising the Firedrake finite element library, in particular how the design of an appropriate computational interface enables these schemes to be used as building blocks inside block preconditioners.

References


We deal with numerical solution of the incompressible Navier–Stokes equations discretized using the isogeometric analysis approach. Isogeometric analysis (IgA) is a discretization approach based on Galerkin method suitable for domains represented as B-spline or NURBS objects, which is usual in industrial practice. IgA has many common features with the finite element method, but instead of creating a computational mesh and using a standard finite element basis, the B-spline/NURBS basis of the geometry representation is used to approximate the solution of the equations. One of the main differences is higher global smoothness of the IgA solutions. Due to the larger supports of the IgA basis functions, the discretization leads to denser matrices than for standard finite elements.

We are interested in iterative solution of the saddle-point type linear systems arising from the IgA discretization of the coupled linearized incompressible Navier–Stokes equations. In this talk we present a comparison of several block preconditioners applied to these systems. We focus on two classes of block preconditioners: block triangular preconditioners based on Schur complement approximation and SIMPLE-type preconditioners.

This is a joint work with C. Vuik (Delft University of Technology, The Netherlands) and J. Egermaier (University of West Bohemia, Czech Republic).

\footnote{University of West Bohemia in Pilsen, Czech Republic (hhornik@kma.zcu.cz)}
The authors extend an (incomplete) LDLT factorization method by Maubach, Lungten and Schilders to the an (incomplete) LDU method for saddle point problems of form $X = [A; B^T; B, 0]$ with non-symmetric positive definite $A$ and $B$ which can be transformed to upper trapezoid form. The LDU (incomplete) factorization is of the form $X = (L + D)D^{-1}(D + U)$ and is such that for symmetric $X$ automatically $U = L$, i.e., there is no longer a difference between the symmetric and non-symmetric cases (an LU factorization does not lead to $U = L^T$ in the symmetric case). Furthermore, for linear problems, specific non-Hermitian complex valued matrices $A$ will topic of examination.

Typical industrial applications are Stokes equations (depending on conditions also Navier-Stokes equations), as well as saddle point systems stemming from linear electrical circuits and for instance quadratic optimization. The LDLT method is now available as prototype C-code, as well as the related LDU code the complex valued case is under construction and has become much faster since its publication. For the LDLT method a MEX interface for Matlab is available. The new implementations are able to combine AMD and other reorderings with the core Schilders reordering, are ensured to converge also with the extra reorderings but do no longer create constraint preconditioners. Numerical results are presented, as well as a comparison with the Matlab built-in ldl method (from Rutherford Laboraties).
Preconditioned iterative methods are now pervasive in performing model reduction of dynamical systems of various orders (first, second, and higher) and differing nonlinearities (linear, bilinear, and nonlinear). The kind of preconditioning required is usually dependent on three characteristics; (a) the order/ nonlinearity of the input dynamical system, (b) the model reduction technique (algorithm) used, and (c) the structure/ sparsity of the input system matrices. This minisymposium focusses on preconditioners that exploit one or more such properties. Further, we also plan to discuss efficient variants of these preconditioning techniques (via updates). Here is the list of speakers with their respective affiliations and titles.

1. **Kapil Ahuja** (IIT Indore): Preconditioned Linear Solves for Parametric Model Order Reduction.

2. **Heidi K. Thornquist** (Sandia National Labs): The Art of Preconditioners for Circuit Simulation.


For computing reduced-order models of parametric dynamical systems, there exist many classes of algorithms commonly termed as parametric model order reduction algorithms. The main computational cost of these algorithms is in solving sequences of very large and sparse linear systems of equations, which are predominantly dependent on slowly varying parameter values. We focus on efficiently solving these linear systems that arise while reducing parametric linear dynamical systems by iterative methods with appropriate preconditioners. We assume that parameters are linearly embedded in the dynamical systems matrices, and specifically work on RPMOR algorithm [1]. We also discuss how our proposed techniques can be applied to other such algorithms.

We propose that the choice of underlying iterative solver is problem dependent. Since for many parametric model order reduction algorithms, the right-hand sides of the linear systems are available together, we propose the use of block variant of the underlying iterative method. Due to constant increase in the input model size and the number of parameters in it, computing a preconditioner in a parallel setting is increasingly becoming a norm. Since Sparse Approximate Inverse (SPAI) preconditioner is a general preconditioner that can be naturally parallelized, we propose its use. Our most novel contribution is a technique to cheaply update the SPAI preconditioner, while solving the parametrically changing linear systems. We support our proposed theory by numerical experiments on two different models. Experimentally, we demonstrate that using a block variant of the underlying iterative solver saves nearly 95% of the computation time over the non-block version. Further, and more importantly, block GCRO with SPAI update saves around 60% of the time over block GCRO with SPAI. More details can be found in [2] (an earlier version of this work).

References


Adaptive Iterative Rational Global Arnoldi (AIRGA) Algorithm [1] is used for model reduction of second order linear dynamical systems with proportional damping. The main computational cost of this algorithm is solving a sequence of linear systems. We propose the use of relevant Krylov subspace method and the Sparse Approximate Inverse (SPAI) preconditioner. Since these linear systems change only slightly from one iteration step to the next, we propose a technique to cheaply update the SPAI preconditioner. Our theory is illustrated by experiments. While reducing a 1.2 million size industrial brake model, we show that on an average, using updates reduces the preconditioner computation time from 10 hrs to 1 hr 40 min (approximately). For more details see [2].

References


For many applications, e.g., parametric model reduction, we need to solve a long sequence or large group of related linear systems. Since a good preconditioner is both essential for fast convergence and expensive to compute, we would like to update preconditioners at low cost. In this presentation, we present methods to update preconditioners, for example, sparse approximate maps or low-rank updates to the identity, that use information that becomes available during the Krylov iteration.
The Art of Preconditioners for Circuit Simulation

H. Thornquist

The ability to perform scalable simulation of very large application-specific integrated circuits (ASICs) is highly dependent upon the availability of effective preconditioners. Since ASICs are a connected network of devices that are designed to perform a wide variety of functions within an electrical system, there is not one single preconditioning strategy that works for all types of ASICs. This talk will discuss the art of designing an effective preconditioner for families of ASICs that generate heterogeneous matrix structures.

1Sandia National Laboratories
Minisymposium # 5: Structured preconditioning

Yuanzhe Xi
Jianlin Xia

In this minisymposium, eight speakers will present their recent work on using broad structure types in preconditioning and iterative solutions. By exploiting structures arising from different problem backgrounds, it is possible to get more reliable and effective iterative solutions. Research results will also be presented on how to apply structured preconditioning to more general situations (such as usual SPD matrices) and challenging problems (such as Helmholtz problems and data science problems).

The minisymposium includes two sessions:

Session 1:

1. Yuanzhe Xi, Emory University
   *Fast contour preconditioner for Helmholtz equations*

2. Kees Vuik, Delft University of Technology
   *A scalable preconditioner for the Helmholtz Equation*

3. Geoffrey Dillon, University of South Carolina
   *Structured preconditioning for fractional PDEs*

4. Mikhail Lepilov, Purdue University
   *Sparse hierarchical preconditioners*

Session 2:

1. Jianlin Xia, Purdue University
   *Effective ESIF preconditioners with guaranteed robustness for general SPD matrices*

2. James Nagy, Emory University
   *Preconditioning of flexible Krylov methods for low rank image reconstruction*

3. Chao Chen, UT Austin
   *A robust hierarchical solver for ill-conditioned systems with applications to ice sheet modeling*

4. Jie Chen, IBM Research
   *Linear-cost covariance functions for Gaussian random fields*

---

1 Emory University
2 Purdue University
Fractional differential equations are an adequate model for many applications such as: anomalous diffusion/dispersion, turbulent flow, porous media flow, quantum mechanics and finance. Unlike partial differential operators, fractional differential operators are nonlocal by nature. As a result, any numerical method for solving these equations has to incorporate this nonlocality, which leads to fully dense linear systems.

Fortunately, these linear systems generally are Toeplitz in structure, and so efficient storage is possible. Perhaps more importantly, we have fast matrix-vector products thanks to the Fast Fourier Transform, making Krylov subspace methods attractive. These are still difficult systems to solve, so preconditioners are required.

Solvers for a 1D fractional order problem with boundary layers were proposed in [1]. In order to resolve these boundary layers, a composite mesh was used that gives rise to a $2 \times 2$ block structured linear system. In this work, we study the effectiveness of various preconditioners for this same problem with a hope of extension to 2D problems in future work.

References


---

1University of South Carolina
2University of South Carolina
In this talk, we propose an iterative solution method for the 3D high-frequency Helmholtz equation. In a contour integration framework, the solution in certain invariant subspace is approximated by solving problems with complex shifts, and this accelerates GMRES iterations by restricting the spectrum. We construct a polynomial fixed-point iteration for solving the shifted problems, which is robust even if the magnitude of the shifts is small. Numerical tests in 3D show that $O(n^{1/3})$ matrix-vector products are needed for solving a high-frequency problem with matrix size $n$ to high accuracy. The method has little storage requirement, can be applied to both dense and sparse linear systems, and is suitable for parallel computing. More details can be found in [1].

References

[1] Xiao Liu, Yuanzhe Xi, Yousef Saad, and Maarten V. de Hoop, Solving the 3D high-frequency Helmholtz equation using contour integration and polynomial preconditioning, arxiv 1811.12378, 2018

1Emory University
2Rice University
3University of Minnesota
4Rice University
Recent research efforts aimed at iteratively solving the Helmholtz equation has focused on incorporating deflation techniques for GMRES-convergence accelerating purposes. The two-level-deflation preconditioner combined with Complex Shifted Laplacian Preconditioner (CSLP) showed encouraging results in moderating the rate at which the eigenvalues approach the origin and cause the solver to slow down. However, for large wave numbers the initial problem resurfaces and the near-zero eigenvalues reappear. Our findings reveal that the reappearance of these near-zero eigenvalues occurs if the near-singular eigenmodes of the fine-grid operator and the coarse-grid operator are not properly aligned. This misalignment is caused by accumulating approximation errors during the inter-grid transfer operations. We propose the use of higher-order approximation schemes to construct the deflation vectors. The results from Rigorous Fourier Analysis (RFA) and numerical experiments confirm that our newly proposed scheme outperforms any deflation-based preconditioner for the Helmholtz problem. In particular, the spectrum of the adjusted preconditioned operator stays fixed near one. For the first time, the convergence properties for very large wavenumbers ($k = 10^6$ in one-dimension, $k = 1.5 \times 10^3$ in two-dimensions and $k = 100$ in three-dimensions) have been studied, and the convergence is almost wave number independent. The new scheme additionally shows very promising results for the more challenging Marmousi problem. For more details, see [1].

References


---

1Delft University of Technology, Department of Numerical Analysis
2Delft University of Technology, Department of Numerical Analysis
Hierarchical SIF Preconditioners for Sparse SPD Matrices

M. Lepilov 1 J. Xia 2 Z. Xin 3

Many sparse preconditioning techniques have been devised over the past few decades. In recent years, preconditioning techniques based on structured incomplete factorization (SIF) have been proposed for general dense SPD matrices and have some attractive features, in particular being amenable to rigorous analysis of their performance. In this talk, preconditioners arising from sparse hierarchical factorization, along with new variants thereof using SIF techniques, will be discussed together with an analysis of their robustness, accuracy, and effectiveness.

This discussion will rely on ideas such as the multifrontal method, nested dissection, and SIF techniques, more details of which may be found in [1], [2], and [3], respectively. This work is joint with Jianlin Xia and Zixing Xin.

References


1 Purdue University
2 Purdue University
3 Purdue University
A hierarchical solver is proposed for solving sparse ill-conditioned linear systems in parallel. The solver is based on a modification of the LoRaSp method [1, 2], but employs a deferred-compression technique, which provably reduces the approximation error and significantly improves efficiency. Moreover, the deferred-compression technique introduces minimal overhead and does not affect parallelism. As a result, the new solver achieves linear computational complexity under mild assumptions and excellent parallel scalability. To demonstrate the performance of the new solver, we focus on applying it to solve sparse linear systems arising from ice sheet modeling. The strong anisotropic phenomena associated with the thin structure of ice sheets creates serious challenges for existing solvers. To address the anisotropy, we additionally developed a customized partitioning scheme for the solver, which captures the strong-coupling direction accurately. In general, the partitioning can be computed algebraically with existing software packages, and thus the new solver is generalizable for solving other sparse linear systems. Our results show that ice sheet problems of about 300 million degrees of freedom have been solved in just a few minutes using a thousand processors, which is much faster than an ILU preconditioner (see figure 2). More details can be found in [3].

![Weak scaling experiment on 1, 4, 16, 64, 256 processors: comparison of the total runtime (factorization+solve) between ILU and our hierarchical solver (hsolver). Dashed line means extrapolation based on existing data because ILU didn’t converge to $10^{-12}$.](image)

---

1 University of Texas at Austin
2 Stanford University
3 Sandia National Laboratories
4 Sandia National Laboratories
5 Sandia National Laboratories
6 Stanford University
References


Gaussian random fields (GRF) are a fundamental stochastic model for spatiotemporal data analysis. An essential ingredient of GRF is the covariance function that characterizes the joint Gaussian distribution of the field. Commonly used covariance functions give rise to fully dense and unstructured covariance matrices, for which required calculations are notoriously expensive to carry out for large data. In this work, we propose a construction of covariance functions that result in matrices with a hierarchical structure. Empowered by matrix algorithms that scale linearly with the matrix dimension, the hierarchical structure is proved to be efficient for a variety of random field computations, including sampling, kriging, and likelihood evaluation. Specifically, with $n$ scattered sites, sampling and likelihood evaluation has an $O(n)$ cost and kriging has an $O(\log n)$ cost after preprocessing, particularly favorable for the kriging of an extremely large number of sites (e.g., predicting on more sites than observed). We demonstrate comprehensive numerical experiments to show the use of the constructed covariance functions and their appealing computation time. Numerical examples on a laptop include simulated data of size up to one million, as well as a climate data product with over two million observations. Code is available at https://github.com/jiechenjiechen/RLCM.

References

In this talk, we show a type of effective and robust preconditioners for general symmetric positive definite (SPD) matrices based on structured incomplete factorization (SIF), called extended SIF (eSIF) preconditioners. The original SIF preconditioning strategy by Xia, et al. applies compression to scaled off-diagonal blocks to yield structured preconditioners. As compared with the original SIF preconditioners, the eSIF preconditioners have three significant advantages:

1. With a similar truncation tolerance in the approximation of certain subblocks, the preconditioner can approximate the original matrix to a much higher accuracy.

2. The preconditioner is much more effective in the sense that the reduction of condition numbers is much more significant, and the eigenvalues of the preconditioned matrix are much better clustered.

3. The multilevel eSIF preconditioners are guaranteed to be positive definite, while the original multilevel SIF preconditioner has a strict requirement in order to preserve positive definiteness.

This is joint work with Mikhail Lepilov.
(MS) Preconditioning of flexible Krylov methods for low rank image reconstruction

James Nagy

Abstract not received
Multigrid methods [1, 2, 3] have been established as being among the most efficient techniques for solving partial differential equations (PDEs). Problems within various disciplines, including fluid dynamics, solid mechanics, geophysics, and more, have been successfully solved by these methods, and recently, some efforts are being carried out for the design of multigrid methods for dealing with problems in many other areas as image processing, optimization, machine learning, and data science. Therefore, the development and analysis of multigrid methods remain an important and fundamental area of research in computational science and engineering. In this minisymposium, we bring together experts focused on different aspects and applications of multigrid methods to discuss the recent advances of multigrid solvers across different disciplines.

In particular, the following speakers accepted to present their recent work in our session:

- Salvatore Filippone - “Efficient algebraic multigrid for scalable scientific simulation”
- Yunhui He - “A local Fourier analysis for additive Vanka relaxation”
- Xiaozhe Hu - “Robust preconditioners for mixed-dimensional models of flow in fractured porous media”
- Ruipeng Li - “Recent Development of Multigrid Solvers in HYPRE on Modern Heterogeneous Computing Platforms”
- Peter Ohm - “Monolithic multigrid for a stabilized discretization of the poroelastic equations”
- Raymond Tuminaro - TBA
- Irad Yavneh - “Accelerating Multigrid via Sequential Subspace Optimization (SESOP)”
- Ludmil Zikatanov - “An auxiliary space preconditioner for mixed finite element discretizations of elliptic equations”

References


1Tufts University (X. Hu) and University of Zaragoza (C. Rodrigo)

Modern many-core processors such as graphics processing units (GPUs) are becoming an integral part of many high performance computing systems nowadays. These processors yield enormous raw processing power in the form of massive SIMD parallelism. Accelerating multigrid methods on GPUs has drawn a lot of research attention in recent years. For instance, in recent releases of the HYPRE package, the structured multigrid solvers (SMG, PFMG) have full GPU-support for both the setup and the solve phases, whereas the algebraic multigrid (AMG) solver, namely BoomerAMG, has only its solve phase been ported and the setup can still be computed on CPUs only. In this talk, we will provide an overview of the available GPU-acceleration in HYPRE and present our current work on performing AMG setup on GPUs. In particular, we focus on the computation of distributed triple-matrix multiplications (i.e., the Galerkin product), which often represents a significant cost of the entire setup phase of AMG. We will discuss in detail the composing parts of this computation that include several fast GPU sparse matrix kernels and communications between GPUs. The recent results as well as future work will also be included.
Multigrid methods are popular solution algorithms for many discretized PDEs, either as standalone iterative solvers or as preconditioners, due to their high efficiency. However, the choice and optimization of multigrid components is crucial to the design of efficient algorithms. It is well-known that local Fourier analysis (LFA) is a useful analysis tool to predict and analyze actual performance of many efficient algorithms for solution of discretized PDEs, especially for the construction of multigrid components, such as relaxation schemes and grid-transfer operators. In this talk, we develop a local Fourier analysis of monolithic multigrid methods based on additive Vanka relaxation schemes for mixed finite-element discretizations of the Stokes equations. Key conclusions of this analysis not only include insight into the choice of “patches” for the Vanka relaxation, but also the comparison of efficiency and effectiveness among several variants. From local Fourier analysis, parameters that minimize the two-grid convergence factor of the multigrid method are proposed. Finally, some numerical experiments are presented to validate the LFA predictions.
Efficient algebraic multigrid for scalable scientific simulation

S. Filippone ¹
P. D’Ambra ²

Algebraic Multigrid variants are popular and effective choices for the preconditioning of Krylov subspace methods in the solution of large and sparse linear systems arising in many application areas.

The race to achieve Exascale computing, that is, the availability of computing platforms capable of executing $10^{18}$ arithmetic operations per second, has produced a substantial amount of research into extreme scalability. Effective preconditioners need to address two separate issues:

**Algorithmic scalability:** the ability to converge in the same number of iterations irrespective of the system size and of the number of computing cores employed;

**Implementation scalability:** the ability to execute computations at the same rate on each computing element, as the size of the system and the number of computing elements grow.

In [¹] we have introduced coarsening software for algebraic multigrid based on Graph Matching techniques in serial mode; an immediate extension by decoupled application was discussed in [²]. These coarsening strategies have proven to be quite effective in enabling algorithmic scalability as the linear system size varies; in this talk we will turn our attention to the use of parallel matching schemes to provide better algorithmic scalability as the computational platform grows to exascale levels. We will discuss the underlying theory, the current implementation status, and will present some experimental results obtained on applications in the context of the Energy Oriented Centre of Excellence EoCoE-II project, funded by the EU.

This work was partially supported by EU Grant 824158 — EoCoE-II.

References


¹ Cranfield University, UK, and IAC-CNR, Italy
² IAC-CNR, Italy
Mixed-dimensional partial differential equations arise in many physical applications including flow in fractured porous media, where the fractures and their intersections form a hierarchy of lower-dimensional submanifolds. An essential component, and usually the most time-consuming part of simulating PDEs, is solving the large-scale and ill-conditioned linear systems of equations arising from discretizations. In this work, we generalize the traditional framework of designing preconditioners for the saddle point systems and develop effective preconditioners that are robust with respect to the physical and discretization parameters for mixed-dimensional models for flow in fractured porous media. Preliminary numerical experiments are presented to support the theory and demonstrate the robustness of our preconditioners.
This talk will introduce a merger of two optimization frameworks: SEquential Subspace Optimization (SESOP) and MultiGrid (MG) optimization. At each iteration of the combined algorithm, search directions implied by the coarse-grid correction process of MG are added to the low dimensional search-spaces of SESOP, which include the preconditioned gradient and search directions involving the previous iterates. The resulting technique is called SESOP-MG. The asymptotic convergence rate of the two-level version of SESOP-MG (dubbed SESOP-TG) is studied via Fourier mode analysis for linear problems, i.e., optimization of quadratic functionals. Numerical tests on linear and nonlinear problems demonstrate the effectiveness of the approach. An evolving manuscript is available in [1].

References

With an ever growing variety of space vehicles, the simulation of hypersonic flight gives rise to increasingly important research challenges associated with the efficient solution of highly non-symmetric systems. While the use of multigrid within sub-sonic and transonic flow simulations has been well studied, the hypersonic flow case has received little attention and is still very much an open research topic.

In this talk, we investigate the application of algebraic multigrid to solve the linear systems that arise from the use of Newton’s method in conjunction with adaptive pseudo-time stepping to solve steady state hypersonic flow problems. The overall approach gives rise to a series of linear systems associated with a Jacobian approximation of the nonlinear equations that is then solved by a multigrid method. Unfortunately, the application of algebraic multigrid can be problematic for highly non-symmetric systems. Specifically, the coarse grid equations produced by a standard Petrov-Galerkin approximation may not be stable, though the fine grid equations are stable. Thus, standard smoothing algorithms such as ILU and line Gauss-Seidel often diverge when applied directly to coarse level equations. Further, solutions from a coarse grid do not necessarily reduce error (as the underlying projections are not orthogonal). We discuss a few algebraic multigrid algorithms to address the above-mentioned concerns. This includes the adaptation of specialized grid transfers that attempt to mimic upwinding ideas used for highly convective flows and the use of non-Galerkin or non-Petrov-Galerkin coarse grid operators. The basic idea behind a non-Petrov-Galerkin approach is to adds a perturbation to the coarse level operator produced by a standard Petrov-Galerkin projection. Two natural possibilities for this matrix perturbation include a diagonal term (motivated by considering a reduced pseudo-time step on the coarse level) as well as a projection portion of the the symmetric part of the matrix (loosely motivated by artificial dissipation considerations). In both cases, the aim is to add the smallest perturbation such that the standard smoothing algorithms will converge on the resulting system.

Numerical results will be given on HPC systems using test problems arising from a hypersonic flow simulation capability, SPARC, that is being developed at Sandia.

---

1Sandia National Laboratories
2Sandia National Laboratories
3Sandia National Laboratories
4Sandia National Laboratories
5Sandia National Laboratories
6University of Wyoming
7University of Wyoming
In this talk, we discuss monolithic multigrid as a preconditioner for a recently developed stabilized discretization of the poroelastic equations [1]. In particular, we consider the three-field formulation of Biot’s consolidation model, which has applications in a wide variety of fields such as geomechanics, petroleum engineering and biomechanics. Previous work on block preconditioners based on the well-posedness of the discrete system yield parameter-robust solution methods [2]. Here, we consider a coupled multigrid approach, applying the preconditioner to the entire system and investigate the parameter robustness in this setting. We use geometric multigrid with finite-element interpolation operators and Galerkin coarsening, focusing on different relaxation schemes. We consider both Braess-Sarazin and an Uzawa-type relaxation, and optimize the relaxation parameters via local Fourier analysis. Results are shown for both the stabilized discretization and a perturbed scheme that allows for the elimination of a significant number of degrees of freedom.

References


We propose an auxiliary space method for the solution of the indefinite problem arising from mixed method finite element discretizations of scalar elliptic problems. The proposed technique uses conforming elements as an auxiliary space and utilizes special interpolation operators for the transfer of residuals and errors between the spaces. We show that the corresponding method provides optimal solver for the indefinite problem by only solving symmetric and positive definite auxiliary problems. We apply this preconditioner to the mixed form discretization of the steady state Richards’ equation linearized with the $L$-scheme. We provide numerical tests validating the theoretical estimates.
While simulation of single-phase fluid flow is already challenging for many reasons, additional challenges arise when considering multiphase or complex fluids that lead to even larger coupled systems of PDEs. The discretization and linearization of these equations often result in algebraic systems with a block structure which can be used to develop efficient and scalable preconditioners. In this minisymposium, recent advances on a range of topics in preconditioning for fluid problems will be presented.

Patrick E. Farrell (University of Oxford): Augmented Lagrangian preconditioners for nematic liquid crystal problems

Massimiliano Ferronato (University of Padova): A class of block preconditioners for the solution of fluid problems in deformable porous media

Robert C. Kirby (Baylor University): Software infrastructure for coupled fluids preconditioners

Scott P. MacLachlan (Memorial University of Newfoundland): Boundary-layer preconditioners for singularly perturbed convection-diffusion equations

Thomas Roy (University of Oxford): Two-stage preconditioners for non-isothermal flow in porous media

Johann Rudi (University of Texas at Austin): An optimization-based perturbed Newton method for viscoplastic fluids with von Mises yielding

John N. Shadid (Sandia National Laboratories): Scalable block preconditioning methods for solution of implicit / IMEX FE continuum plasma physics models

Martin Stoll (TU Chemnitz): Low-rank solvers and preconditioning for unsteady Stokes-Brinkman optimal control problem with random data
A class of block preconditioners for the solution of fluid flow problems in deformable porous media

M. Ferronato

M. Frigo

N. Castelletto

MS: Preconditioners for Fluid Problems

Fully implicit numerical models simulating fluid flow in deformable porous media are based on non-linear poromechanics theory and require the solution of large size systems of equations with a block structure. In this work, we focus on a class of block preconditioners for the linearized $3 \times 3$ block system arising from the mixed finite element discretization of a three-field formulation for coupled poromechanics. The preconditioner is obtained by: (i) combining proper physics-based splittings of the block matrix, and (ii) introducing an optimal relaxation parameter $\alpha$, with the objective to avoid the computation of approximate Schur complements. Similar developments were first proposed in [1, 2] for the saddle-point problems arising from Navier-Stokes equations. A stable and cost-effective algorithm is here advanced to compute the optimal value of $\alpha$. However, a possible drawback relies on the need of inverting inner blocks in the form $\hat{C} = C + \frac{1}{\alpha} FF^T$, where $C$ is a regular square matrix and $FF^T$ is largely rank-deficient, for small values of $\alpha$. Novel strategies are introduced for the efficient and stable approximate application of $\hat{C}^{-1}$ to a vector based on the use of a projection operator onto the range of $F$. The proposed class of block preconditioners is tested in both theoretical benchmarks and real-world large-size applications, proving an effective and robust alternative to other approaches advanced in the literature.

Portions of this work were performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

References


1Department ICEA, University of Padova, Padova, Italy
2Department ICEA, University of Padova, Padova, Italy
3Atmospheric, Earth and Energy Division, Lawrence Livermore National Laboratory, Livermore, CA, USA
Embedded in many algorithms for the simulation of problems in computational fluid dynamics is the need to accurately and efficiently model boundary layers in the flow. In this talk, we consider the simpler model of a scalar convection-diffusion equation, where the structure of the arising boundary layers is well-known in existing theory. This structure motivates the use of layer-adapted meshes on which standard finite-difference discretizations give approximation of the solution that is independent of the layer widths: so-called “parameter-robust methods”. We propose preconditioners that also reflect the layer structure for convection-diffusion equations and, consequently, are also robust with respect to layer width. In this talk, we will present theory for the one-dimensional case, along with heuristics and numerical results for the two-dimensional case.

1Department of Mathematics and Statistics, Memorial University of Newfoundland
2School of Mathematics, Statistics, and Applied Mathematics, National University of Ireland, Galway
3Department of Mathematics and Science, Holy Names University
(MS) Preconditioners for Fluid Problems: Software infrastructure for coupled fluids preconditions

Robert C. Kirby
Lawrence Mitchell

The efficient solution of discretizations of coupled systems of PDE is at the core of much numerical simulation. Significant effort has been expended on scalable algorithms to precondition Krylov iterations for the linear systems that arise. With few exceptions, the reported numerical implementation of such solution strategies is specific to a particular model setup, and intimately ties the solver strategy to the discretization and PDE, especially when the preconditioner requires auxiliary operators. In this talk, we present recent improvements in the Firedrake finite element library that allow for straightforward development of the building blocks of extensible, composable preconditioners that decouple the solver from the model formulation. Our implementation extends the algebraic composability of linear solvers offered by the PETSc library by augmenting operators, and hence preconditioners, with the ability to provide any necessary auxiliary operators. Rather than specifying up front the full solver configuration tied to the model, solvers can be developed independently of model formulation and configured at runtime. We illustrate with examples incompressible fluids, temperature-driven convection, and tide modeling.

This talk is based on work published in [?] and unpublished work on preconditioning the model analyzed in [?]

References


1Baylor University
2Durham University
In this talk, we propose a preconditioner of augmented Lagrangian type for the Oseen–Frank free energy model for nematic liquid crystal problems. The augmented Lagrangian term achieves control of the Schur complement of the block system derived after Newton linearization and spatial discretization, but makes the solution of the augmented $(1,1)$ block more challenging. The approach is motivated by its success in devising Reynolds-robust preconditioners for the Navier–Stokes equations [1, 2], where the linear incompressibility constraint must be enforced. A key difference in the Oseen–Frank model is that the constraint to be enforced is nonlinear, yielding additional challenges for the solution of the augmented problem.

Following Schöberl’s analysis for constructing parameter-robust multigrid schemes [3], we will consider two essential ingredients: kernel-preserving relaxation and parameter-robust prolongation. We will show that for certain discretisation choices, Schöberl’s conditions can be satisfied to devise a robust multigrid solver for the augmented $(1,1)$ block, yielding parameter-robust convergence for the entire preconditioner for the Oseen–Frank problem. Numerical experiments are presented to demonstrate the results.

References


The mathematical basis for the continuum modeling of plasma physics systems is the solution of the governing partial differential equations (PDEs) describing conservation of mass, momentum, and energy, along with various forms of approximations to Maxwell’s equations. The resulting systems are characterized by strong nonlinear and nonsymmetric coupling of fluid and electromagnetic phenomena, as well as the significant range of time- and length-scales that these interactions produce. To enable accurate and stable approximation of these systems a range of spatial and temporal discretization methods are commonly employed. In the context of finite element spatial discretization methods these include mixed integration, stabilized and variational multiscale (VMS) methods, and structure-preserving (physics compatible) approaches.

For effective long-time-scale integration of these systems the implicit representation of at least a subset of the operators is required.

Two well-structured approaches, of recent interest, are fully-implicit and implicit-explicit (IMEX) type time-integration methods employing Newton-Krylov type nonlinear/linear iterative solvers. To enable robust, scalable and efficient solution of the large-scale sparse linear systems generated by a Newton linearization, fully-coupled multilevel preconditioners are developed. The multilevel preconditioners are based on two differing approaches. The first technique employs a graph-based aggregation method applied to the nonzero block structure of the Jacobian matrix. The second approach utilizes approximate block factorization (ABF) methods and physics-based preconditioning approaches that reduce the coupled systems into a set of simplified systems to which multilevel methods are applied.

To demonstrate the flexibility of implicit/IMEX FE discretizations and the fully-coupled Newton-Krylov-AMG solution approaches various forms of resistive magnetohydrodynamic (MHD) and multifluid electromagnetic plasma models are considered. In this context, we first briefly discuss the mathematical models and formulations for a subset of these systems, and then present results for representative plasma physics problems of current scientific interest. Additionally, the discussion considers the robustness, efficiency, and the parallel and algorithmic scaling of the preconditioning methods. Weak scaling results include studies on up to 1M cores. (This is joint work with Roger Pawlowski, Edward Phillips, Paul Lin, Sidafa Conde, Eric Cyr, Sean Miller and Sibu Mabuza)

*This work was partially supported by the DOE Office of Science Advanced Scientific Computing Research (ASCR) - Applied Math Research program and an ASCR/Office of Fusion Energy SciDAC Partnership Project at Sandia National Laboratories.

1Sandia National Laboratories, Center for Computing Research, Albuquerque NM; Department of Mathematics and Statistics, University of New Mexico
An Optimization-Based Perturbed Newton Method for Viscoplastic Fluids with von Mises Yielding

Johann Rudi
Georg Stadler
Omar Ghattas

We target highly nonlinear fluid models with von Mises yield criteria. Due to the yield criteria, an initially viscous fluid model switches to a plasticity model when the second deviatoric stress invariant reaches a critical value. Von Mises yielding occurs, e.g., in nonlinear incompressible Stokes flow in Earth’s mantle with plastic yielding rheology \cite{3}, where the deviatoric stress is limited by the critical value called yield strength. Therefore, if the strain rates increase, the effective viscosity decreases such that the maximum stresses are bound by this yield strength value.

Such a switch between viscous and plastic models, which is incorporated into the constitutive relationship between strain rate and deviatoric stress, gives rise to an optimization problem where the Hessian exhibits a (near) null space upon linearization with Newton’s method. The null space of the Hessian operator is caused by a projector-type coefficient in the Hessian, which arises from a term in the objective functional that resembles the $L^1$-norm. Using a standard Newton linearization is known to produce severe Newton step length reductions due to backtracking line search and stagnating nonlinear convergence. Additionally, these effects become increasingly prevalent as the mesh is refined. We consider the null space of the Hessian operator to cause the problematic convergence of Newton’s method.

In this presentation, we analyze issues with the standard Newton linearization in an abstract setting and propose an improved linearization, which can be applied straightforwardly to Stokes flow with yielding. Our improved Newton linearization is motivated by and generalizes results from image restoration applications that use total variation regularization \cite{1, 2}, which is formed by the $L^1$-norm of the gradient.

Numerical experiments compare the standard and improved Newton linearizations for mantle flow applications with von Mises yielding. When we employ our improved linearization within our inexact Newton–Krylov method, a fast and highly robust nonlinear solver is attained that exhibits (nearly) mesh-independent convergence and, in addition, scales to large numbers of cores with high parallel efficiency.

---

1. Mathematics and Computer Science Division, Argonne National Laboratory
2. Courant Institute of Mathematical Sciences, New York University
3. Oden Institute, Jackson School of Geosciences, and Department of Mechanical Engineering, The University of Texas at Austin
References


We consider the simulation of an optimal control problem constrained by the unsteady Stokes-Brinkman equation involving random data. We consider a generalized polynomial chaos approximation of these random functions in the stochastic Galerkin finite element method. The discrete problem yields a prohibitively high dimensional saddle point system with Kronecker product structure. We discuss how such systems can be solved using tensor-based methods and efficient preconditioning strategies. The performance of our approach is illustrated with extensive numerical experiments.
In petroleum reservoir simulation, the standard preconditioner, Constrained Pressure Residual (CPR) [1], is a two-stage process which involves solving a restricted pressure system. Initially designed for isothermal models, this approach is often used in the thermal case. However, its treatment of the temperature variable does not incorporate heat diffusion, which is often dominant in thermal cases. In this talk, we present preconditioners which consider heat diffusion as well as the cross-coupling between pressure and temperature. In order to study the effects of both pressure and temperature on fluid and heat flow, we first consider a model of non-isothermal single phase flow in porous media. For this model, we develop a block preconditioner with an efficient Schur complement approximation [2]. Then, we extend this method for multiphase flow as a CPR-like two-stage preconditioner. The algorithmic performance of various two-stage preconditioners is evaluated for reservoir simulation test cases.

References


Complex Polynomial Preconditioners for Indefinite Systems

X. Ye \textsuperscript{1}
Y. Saad \textsuperscript{2}
Y. Xi \textsuperscript{3}

A polynomial preconditioner for solving large sparse indefinite linear systems will be presented. The advantage of such a polynomial form is that only matvec is needed and there is no inner product when applying the preconditioner, which is beneficial for parallel environment and GPUs. An Arnoldi-like algorithm called Proxy-GMRES is proposed to efficiently construct such a polynomial function, the operations are done in a carefully designed polynomial space and some error analysis will be discussed. In addition, we will talk about several improvements for more efficiency and numerical stability.
A New Stable Polynomial Preconditioned GMRES

R. Morgan \textsuperscript{1}

J. Loe \textsuperscript{2}

Polynomial preconditioning can be remarkably effective for difficult systems of linear equations. We give a new polynomial preconditioner for GMRES. A “GMRES polynomial” is used to precondition GMRES, and it is fairly simple to implement even with a complex spectrum. This can be made stable for high degree polynomials. This approach can significantly reduce the number of matrix-vector products and dramatically reduce vector operations such as dot products. Polynomial preconditioning can be combined with regular preconditioning. We will also address some of these topics: how doubling up on roots can make the polynomial more stable, comparisons to BiCGStab and FGMRES, potential performance using Chebyshev estimates, need for a random starting vector to generate the polynomial, implementation of the “other” polynomial that is needed, double polynomial preconditioning, and issues for indefinite problems.

\textsuperscript{1}Baylor University

\textsuperscript{2}Baylor University
Since polynomial preconditioners reduce orthogonalization expense in GMRES, they can be useful for avoiding communication in parallel computing. In particular, the GMRES polynomial can be a very effective general preconditioner for nonsymmetric matrices. We discuss implementation of the preconditioner in the Trilinos package Belos and demonstrate options for maintaining stability. We show several parallel examples of cost reduction.
L-Sweeps: A scalable parallel preconditioner for the high-frequency Helmholtz equation

M. Taus ¹
L. Zepeda-Núñez ²
R. J. Hewett ³
L. Demanet ⁴

In many science and engineering applications, solving time-harmonic high-frequency wave propagation problems quickly and accurately is of paramount importance. For example, in geophysics, particularly in oil exploration, such problems can be the forward problem in an iterative process for solving the inverse problem of subsurface inversion. It is important to solve these wave propagation problems accurately in order to efficiently obtain meaningful solutions of the inverse problems: low order forward modeling can hinder convergence. Additionally, due to the volume of data and the iterative nature of most optimization algorithms, the forward problem must be solved many times. Therefore, a fast solver is necessary to make solving the inverse problem feasible. For time-harmonic high-frequency wave propagation, obtaining both speed and accuracy is historically challenging.

Recently, there have been many advances in the development of fast solvers for such problems, including methods which have linear complexity with respect to the number of degrees of freedom. While most methods scale optimally only in the context of low-order discretizations and smooth wave speed distributions, the method of polarized traces [1] has been shown to retain optimal scaling for high-order discretizations, such as hybridizable discontinuous Galerkin methods and for highly heterogeneous (and even discontinuous) wave speeds [4]. To date, this method relies on a layered domain decomposition together with a preconditioner applied in a sweeping fashion, which has limited straightforward parallelization.

In this work, we introduce a new preconditioning technique for the high-frequency Helmholtz equation based on a checkerboard domain decomposition. The novelty of the new technique is that it can be applied in parallel and is independent of the discretization. In particular, the preconditioner can be applied in $O(N/p)$ complexity where $N$ is the number of degrees of freedom and $p = O(N^{1/d})$ is the number of processors. We consider several numerical examples including constant and non-constant wave speed distributions. Using a preconditioned GMRES method, in all considered examples, the preconditioner results in a logarithmic growth of the number of iterations with respect to the frequency $\omega$. Up to logarithmic factors, this results in the optimal parallel complexity of $O(N/p)$ to solve the resulting linear system. To the best of our knowledge, this is the first such result for the high-frequency Helmholtz equation.

¹TU Wien, Massachusetts Institute of Technology
²Lawrence Berkeley National Laboratory
³Virginia Tech
⁴Massachusetts Institute of Technology

74
References


An Approximate Deflation Preconditioning Method Based on Multiple Grids for Wave Scattering Problems.

Josef Sifuentes
Jesus Saldaña

In this talk I will address adaptive preconditioning algorithms for discretizations of the Lippmann-Schwinger (LS) integral equations. There are two main methods for modeling inhomogeneous wave scattering problem: (1) Helmholtz Partial Differential Equations and (2) (LS) integral equations. While solutions of the (LS) equations automatically satisfy outgoing boundary conditions, the resulting coefficient matrices are large, dense, and non-symmetric. Thus any preconditioning algorithms must play nice with fast matrix vector algorithms (e.g. Fast Fourier Transforms or Fast Multipole Methods). We present an adaptive deflation method that builds approximate invariant subspaces via the Arnoldi iterations already built into GMRES. Furthermore we show that for some inhomogeneities, deflating low frequency eigenvectors substantially improves GMRES convergence. We demonstrate then that the approximate invariant subspaces can be computed on coarse meshes. However care must be taken when interpolating back onto the fine grid that orthogonality is maintained. We will share early work that demonstrates that this method is more generally applicable to problems for which multigrid approaches have been useful. Additionally, I will present convergence bounds for approximations of preconditioners for which the convergence behavior is known.

1University of Texas Rio Grande Valley
2University of Texas Rio Grande Valley
Preconditioned global Krylov subspace methods for solving saddle point systems with multiple right-hand sides

A. Badahmane ¹
A. Bentbib ²
H. Sadok ³

In the present paper, we propose a preconditioned global approach as a new strategy to solve multiple linear systems with several right-hand sides coming from saddle point problems. The preconditioner is obtained by replacing some $(2, 2)$ block in the original Saddle-point matrix $\mathcal{A}$ by another well chosen block. We applied the global GMRES method to solve this new problem with several right hand sides, we gave some convergence results of the preconditioned global GMRES and MINRES. We analyzed the eigenvalue-distribution and the eigenvectors of the proposed preconditioner. Finally, numerical results show that our preconditioned global GMRES method, has high performance as compared to other preconditioned global GMRES methods for solving the saddle point problem with several right hand sides.

References


¹FST-Marrakech, Laboratoire de Mathématiques Appliquées et Informatique, Marrakech, Morocco. LMPA, Université du Littoral Côte d’Opale, 50 Rue F. Buisson, BP 699 - 62228 Calais cedex, France, email: badahmane.achraf@gmail.com.

²FST-Marrakech, Laboratoire de Mathématiques Appliquées et Informatique, Marrakech, Morocco, email: a.bentbib@uca.ac.ma.

³LMPA, Université du Littoral Côte d’Opale, 50 Rue F. Buisson, BP 699 - 62228 Calais cedex, France, email: Hassane.sadok@univ.littoral.fr
Preconditioners for Double Saddle Point Systems

S. Bradley

This talk considers preconditioners for block linear systems of the form:

\[
\begin{bmatrix}
A & B^T & 0 \\
B & -D & C^T \\
0 & C & E
\end{bmatrix}
\begin{bmatrix}
x \\
y \\
z
\end{bmatrix}
= \begin{bmatrix}
f \\
g \\
h
\end{bmatrix},
\]

where $A$, $D$, and $E$ are positive semidefinite, with $D$ and $E$ possibly equal to zero. We refer to these as double saddle point systems. These systems arise in various applications including PDE-constrained optimization, interior point methods, and various formulations and methods for the numerical solution of PDEs.

In this talk we apply Schur complement-based methods to a model problem in PDE-constrained optimization. By exploiting the double saddle point structure and using some new techniques to approximate the Schur complements for this problem, we obtain new preconditioners that are competitive with current state-of-the-art approaches.

\footnote{Department of Computer Science, University of British Columbia}
Dimension-wise splitting iteration with selective relaxation for saddle point problems

Y. Xu 1
M. J. Gander 2
Q. Niu 3
H. Zhang 4

This talk will address a novel relaxed iterative method for the incompressible Navier-Stokes equations and its extension to the optimal control problems, particularly on the choice of the relaxation parameter, more details can be found in [1, 2].

We propose a new Dimension-wise Splitting with Selective Relaxation (DSSR) method for saddle point systems arising from the discretization of incompressible Navier-Stokes equations. When applied to saddle point systems arising from the discretization of optimal control problem, the idea leads to a State Control Alternating Iterative Method (SCAIM). Using Fourier analysis, we determine the optimal choice of the relaxation parameter that leads to the best performance of the iterative method, for the Stokes and the steady Oseen equations for DSSR, and for the optimal control problem governed by elliptic PDEs for SCAIM. We explore numerically the influence of boundary conditions on the optimal choice of the parameter, and the use of inner and outer iterations.

References


We consider solving system of nonlinear algebraic equations arising from the discretization of partial differential equations from computational fluid or solid mechanics. When the nonlinearities in the system is well-balanced, Newton’s method works well, but when a small number of nonlinear functions in the system are much more nonlinear than the others, Newton may converge slowly or even stagnate. In such a situation, we introduce some nonlinear preconditioners to balance the nonlinearities in the system. For the nonlinearly preconditioned problem, we show that fast convergence can be restored. In this talk we discuss some recent progress in the applications of nonlinear preconditioners for some difficulty problems arising in computational biomechanics including both fluid dynamics and solid mechanics.
The estimation of large sparse inverse covariance matrices is an ubiquitous statistical problem in many application areas such as mathematical finance or geology or many others. Numerical approaches typically rely on the maximum likelihood estimation or its negative log-likelihood function. When the Gaussian mean random field is expected to be sparse, regularization techniques which add a sparsity prior have become popular to address this issue. Recently a quadratic approximate inverse covariance method (QUIC) [1] has been proposed. The hallmark of this method is its superlinear to quadratic convergence which makes this algorithm to be among the most competitive methods. In this talk we present a sparse version (SQUIC) [2] of this method and we will demonstrate that using advanced sparse matrix technology the sparse version of QUIC is easily able to deal with problems of size one million within a few minutes on modern multicore computers. We will further address the use of incomplete factorization methods [3] to deal with core parts of the maximum likelihood process such as the computation of the determinant or inverting the precision matrix and also demonstrate how the optimization part can be accelerated using block coordinate descent updates.

References


1 TU Braunschweig
2 USI Lugano
3 USI Lugano
Preconditioning for Toeplitz systems has been well studied over the past few decades. For (real) symmetric Toeplitz systems, descriptive convergence bounds of the preconditioned conjugate gradient method are well understood. Yet, as for nonsymmetric Toeplitz systems, convergence theory is less developed and most work in the literature has been focused on their normal equation systems. In this work, without normalizing a given nonsymmetric Toeplitz system, we symmetrize it by using a simple permutation matrix in order to employ the minimal residual method with theoretical convergence guarantees. In the ill-conditioned case where circulant preconditioners are unsatisfactory, we propose several band Toeplitz preconditioners and show that superlinear convergence can be achieved. Numerical examples are provided to illustrate the effectiveness of our proposed preconditioning strategy.
Low-synch Gram-Schmidt projection schemes applied to GMRES-AMG moving mesh solvers

K. Świrydowicz ¹
S. Thomas ²
T. Warburton ³ A. Austin ⁴ S. Ananthan ⁵

We apply the low-synch classical Gram-Schmidt algorithm recently derived in [1] to the least-squares polynomial projection schemes introduced by P. Fischer in [2]. The original projection algorithm is modified to instead compute a least-squares initial guess for a preconditioned GMRES-AMG pressure solver for the incompressible Navier-Stokes equations. A correction step is first applied to a projected vector perpendicular to the span of previous solutions. An initial guess is formed from the perpendicular and parallel spaces, analogous to a coarse-grid correction. For the Nalu-Wind model, the number of GMRES-AMG iterations is reduced by 6x when applied to moving mesh simulations for rotating wind-turbines. The GMRES iteration count drops to zero with a relatively small window of previous solutions.

References


¹National Renewable Energy Laboratory
²National Renewable Energy Laboratory
³Virginia Tech
⁴Virginia Tech
⁵National Renewable Energy Laboratory
An HSS Preconditioner for Evolving Kernel Matrix Systems

X. Xing

E. Chow

In Stokesian [1] and Brownian [2] dynamics with hydrodynamic interactions, long-range interactions between particles are described by a non-compact kernel function, leading to a dense kernel matrix defined by particle positions. Such a kernel matrix changes along with the evolving particle positions at different time steps in the numerical simulation. In Stokesian dynamics, both matrix-vector multiplications and solves of this kernel matrix are needed at each time step which dominate the computational cost. In this presentation, we illustrate a preconditioner in a hierarchically semiseparable (HSS) [3] matrix form for the iterative solve of these kernel matrices. The key feature is that this HSS preconditioner can be efficiently updated when the kernel matrix evolves along with the simulation.

The kernel-independent fast multipole method (KIFMM) [4] is commonly used to accelerate matrix-vector multiplications of these evolving kernel matrices with linear complexity, which also facilitates iterative solves of the matrices. Meanwhile, fast direct solve algorithms, such as HSS matrices with the ULV decomposition [3], recursive skeletonization (RS) [5], hierarchical interpolative factorization (HIF) [6], and the inverse fast multipole method (IFMM) [7], can approximately factorize and then solve a kernel matrix with asymptotically subquadratic complexity in many cases. In these algorithms, the approximate factorization step can take multiple orders of magnitude more time than the solve step, while the solve step can be as fast as the application of KIFMM. Thus, these algorithms are ideally suited for linear systems with multiple right-hand sides but not for those with evolving kernel matrices. Alternatively, these algorithms can also be used as preconditioners for iterative solves by approximately factorizing the kernel matrix with low accuracy, which has reduced but still expensive computational cost compared with the direct solve by a more accurate approximate factorization.

To efficiently solve evolving kernel matrices, it is natural to use an iterative solver while applying KIFMM to accelerate matrix-vector multiplications and applying a preconditioner to accelerate iteration convergence. Inspired by the RS algorithm which is a specific way of constructing and then factorizing an HSS approximation of a kernel matrix, we choose to construct a preconditioner in an HSS form that approximates the kernel matrix with low accuracy. To allow for efficient update of this preconditioner, we design and apply a special HSS form for the preconditioner with an “equivalent charge” idea borrowed from KIFMM. With this new HSS form, updating the preconditioner only requires a small fraction of the HSS construction step in RS and a full HSS factorization (see the ULV decomposition in Ref. [3]). For an HSS matrix, the HSS factorization is much cheaper than the HSS construction used in RS. As a result, the preconditioner update can be much more efficient than directly applying RS from scratch.

1School of Mathematics, Georgia Institute of Technology
2School of Computational Science and Engineering, Georgia Institute of Technology
References


SpaND: An Algebraic Sparsified Nested Dissection Algorithm Using Low-Rank Approximations

E.G. Boman 1
L. Cambier 2
C. Chen 3
E. Darve 4
S. Rajamanickam 5
R. Tuminaro 6

We propose a new algorithm for the fast solution of large, sparse, symmetric positive-definite linear systems, SpaND — Sparsified Nested Dissection. It is based on nested dissection, sparsification and low-rank compression. After eliminating all interiors at a given level of the elimination tree, the algorithm sparsifies all separators corresponding to the interiors. This operation reduces the size of the separators by eliminating some degrees of freedom but without introducing any fill-in. This is done at the expense of a small and controllable approximation error. The result is an approximate factorization that can be used as an efficient preconditioner. We then perform several numerical experiments to evaluate this algorithm. We demonstrate that a version using orthogonal factorization and block-diagonal scaling takes less CG iterations to converge than previous similar algorithms on various kinds of problems. Furthermore, this algorithm is provably guaranteed to never break down and the matrix stays symmetric positive-definite throughout the process. We evaluate the algorithm on some large problems show it exhibits near-linear scaling. The factorization time is roughly $O(N)$ and the number of iterations grows slowly with problem size $N$. SpaND [1] can be viewed as an extension of the HIF method [2].

References


1 Sandia National Labs. Sandia is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energys National Nuclear Security Administration under contract DE-NA-0003525.
2 Stanford University.
3 Stanford University.
4 Stanford University.
5 Sandia National Labs.
6 Sandia National Labs.
High-performance Deflated Conjugate Gradient Method

J. Kruzik

D. Horak

The talk deals with the choice of the deflation space for the deflated conjugate gradients (DCG) as well as with the DCG coarse problem solution employing a multilevel approach. A PETSc-based implementation of the method will demonstrate the effectiveness of the method on various large-scale benchmarks.

The conjugate gradient (CG) algorithm is often the method of choice for the solution of large symmetric positive definite linear systems of the form

$$Ax = b.$$ 

In order to accelerate the convergence of CG we often need a suitable preconditioner. However, there also exists a complementary approach to the preconditioning known as deflation. The deflation utilizes a deflation space that should represent slowly converging components of the solution.

The deflated conjugate gradient (DCG) method \cite{1}, introduced in \cite{2, 3, 4}, works by splitting the solution of the linear system into two parts. The first part represents the solution in the deflation space and is directly obtained. The second one is computed by CG iterations that operate only on the $A$-conjugate complement of the deflation space.

Given a full rank deflation matrix $W$ whose columns span the deflation space, we can create a projection on the $A$-conjugate complement of $W$

$$P = I - W (W^T AW)^{-1} W^T A.$$ 

This projection is then used inside CG iterations to keep the solution (and the descend directions) in the $A$-conjugate complement of the deflation space.

It can be shown \cite{5} that DCG act as CG “preconditioned” by the projector $P$ as the convergence is governed by the spectrum of $PA$ operator. Moreover, $PA$ operator with a good choice of the deflation space has some of its eigenvalues shifted to zero (deflated).

A good choice of deflation space is crucial for making DCG converge quickly. In practice, there were two main deflation spaces.

The first one uses eigenvectors of $A$ as the deflation space. The associated eigenvalues of the eigenvectors belonging to the deflation space are shifted to zero in the spectrum of the DCG operator $PA$. Particularly, eigenvectors belonging to the smallest eigenvalues are used as they

\footnote{1Institute of Geonics of the Czech Academy of Sciences & VSB-Technical University of Ostrava}

\footnote{2Institute of Geonics of the Czech Academy of Sciences & VSB-Technical University of Ostrava}
slow down the convergence of CG the most. In our experiments, this approach works very well. The problem is how to obtain the eigenvectors.

The second approach is subdomains aggregation. Given a decomposition of the computational domain, each subdomain contributes a single vector into the deflation space. This vector contains ones on the indices of unknowns belonging to the subdomain and zeros otherwise. Such space often approximates a similar space as in the eigenvector approach. We can use, e.g., METIS to obtain the domain decomposition. However, assuming a single computational core owns the whole subdomain then, to utilize the cores appropriately, the subdomains have to be fairly large making the deflation space too coarse to be effective.

A new approach based on wavelet compression was suggested in [6]. The basic idea is that given the wavelet scaling coefficients \(h_1, \ldots, h_k\) we create a projection onto the scaling subspace

\[
H_{1,n} = \begin{pmatrix}
h_1 & h_2 & h_3 & \cdots & 0 & \cdots & 0 & 0 \\
0 & 0 & h_1 & h_2 & \cdots & 0 & 0 & \vdots \\
h_{k-1} & h_k & 0 & 0 & \cdots & h_{k-3} & h_{k-2} & \vdots \\
\end{pmatrix} \in \mathbb{R}^{2^n \times n}.
\]

Then \(H_{1,n}A H_{1,n}^T\) contains trends of \(A\). Moreover, we can repeat this compression process to use up to \(m\) levels of the compression

\[
H_{1,n/2^{m-1}} \cdots H_{1,n/2} H_{1,n} A H_{1,n}^T H_{1,n/2}^T \cdots H_{1,n/2^{m-1}} = H_{m,n} A H_{m,n}^T
\]

Since \(H_{m,n}\) cuts off the high frequencies, we can set \(W = H_{m,n}^T\).

The suggested wavelet compression is also used in the algebraic multigrid [7]. Therefore, using the prolongation matrices from multigrid in place of the deflation matrix might work as well. Moreover, the prolongation operators can be chained, as in the wavelet-based deflation, without the use of any smoothers between multigrid levels.

While large deflation spaces can be highly effective in decreasing the number of iterations, they also make the solution of the inverse (coarse problem) in the projector \(P\) difficult. To alleviate this problem, we turn to the multilevel deflation [8], where DGG is recursively used to solve the coarse problem on each level until the coarse problem is sufficiently small to be quickly solved by a direct solver.

In order to evaluate the aforementioned deflation spaces, an efficient, parallel implementation of DCG was created. It is written as a solver for linear systems in PETSc [9] (KSP). Currently, it is part of the PETSc-based, open-source PERMON library [10].

The benchmarks used in the numerical experiments include, e.g., matrices from SuiteSparse Matrix collection, 3D linear elasticity multi-material cantilever beam and 2D Laplace discretized by boundary element method on an L-shaped domain. Some of the results with appropriate discussion can be found in [5].
References


GeMSLR: A Multilevel Low-Rank Preconditioning and Solution Package

T. Xu \(^1\)
V. Kalantzis \(^2\)
G. Dillon \(^3\)
Y. Xi \(^4\)
R. Li \(^5\)
Y. Saad \(^6\)

This talk summarizes the development and implementation of GeMSLR (Generalized Multilevel Schur complement Low-Rank), a distributed-memory preconditioner for the solution of large and sparse (non)symmetric linear systems of equations. The GeMSLR preconditioner is purely algebraic and is based on a multilevel reordering of the original set of equations/variables \(^1\)~\(^2\). This reordering is implemented by hierarchically ordering the interface degrees of freedom at each level and several reordering schemes are available (Nested Dissection, Independent Sets Ordering, Multicoloring, and Recursive K-way).

At each given level, GeMSLR decouples the solution of the current linear system into one associated with the interior variables and another associated with the interface ones. The first subproblem is block-diagonal and solved in parallel by applying some form of ILU preconditioning. The recursive nature of the preconditioner appears on the second subproblem where the Schur complement linear system is preconditioned by the interface coupling matrix. The latter is applied by descending to the next level until the last level is reached. In the latter case, the user can choose to use either Block Jacobi acceleration or redundantly solve the problem by (I)LU. Low-rank correction terms can be added at each level to further enhance robustness, and these are applied using the Woodbury formula.

GeMSLR is implemented in MPI and has been tested on E5-2680v3 processors in Minnesota Supercomputing Institute’s Mesabi cluster. In addition, GeMSLR can take advantage of nodes equipped with hardware accelerators (e.g. GPU) for certain dense kernels. We demonstrate the potential of GeMSLR by presenting numerical tests performed on several 2D and 3D problems, and both strong and weak scaling is discussed.

---

\(^1\)Department of Computer Science and Engineering, University of Minnesota
\(^2\)Thomas J. Watson Research Center, IBM Research
\(^3\)Department of Mathematics, University of South Carolina
\(^4\)Department of Mathematics, Emory University
\(^5\)Center for Applied Scientific Computing, Lawrence Livermore National Laboratory
\(^6\)Department of Computer Science and Engineering, University of Minnesota
References


Preconditioning sparse SPD linear systems with multiple right-hand sides by recycling and reverse Galerkin projections

V. Kalantzis

This talk discusses subspace recycling preconditioners combined with domain decomposition to accelerate the algebraic solution of sparse symmetric positive-definite linear systems with multiple right-hand sides. Domain decomposition decouples the solution for each right-hand side into two separate subproblems involving the solution of multiple right-hand sides: a) one associated with the interior variables, and b) one associated with the interface variables. Our main focus lies on the solution of the Schur complement multiple right-hand sides problem. We present a deflation-based preconditioner which exploits Krylov subspace recycling and Galerkin projections [1, 2]. In particular, the proposed approach retains the Krylov subspace generated by solving for one or a few right-hand sides, and uses this subspace to accelerate the solution of subsequent solves by generating initial guesses which are (approximately) orthogonal to the captured Krylov subspace. The Galerkin projections are performed in a reverse manner so as to reduce the effects of finite-precision arithmetic.

The proposed preconditioner has several advantages: a) it is applied on the Schur complement which typically has a lower effective condition number than the original iteration matrix, b) it involves no orthogonalization, c) it is suitable as an out-of-core preconditioner. Combinations with algebraic recursive multilevel solvers are possible, and several practical aspects are discussed. Experiments performed on both general SPD matrices and matrices arising from applications in eigenvalue computations demonstrate the competitiveness of the proposed preconditioner against deflated Conjugate Gradient and eigCG [3].

References


1IBM Research, Thomas J. Watson Research Center, Yorktown Heights, NY 10598, USA
Preconditioning techniques for a class of block three-by-three matrices

F. P. A. Beik
M. Benzi

Consider the following system of equations

\[ \begin{bmatrix} A_{11} & A_{12} & 0 \\ A_{21} & A_{22} & B^T \\ 0 & B & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = b, \quad (11) \]

where \( A_{11}, A_{22} \) are both symmetric positive definite matrices, \( A_{21} = -A_{12}^T \) and \( B^T \) is a full column rank matrix. Linear systems of the form (11) arise, for instance, from finite element discretizations of coupled Stokes-Darcy flow [2, 3]. Liquid crystal modeling [5] also leads to sequences of linear systems that can be brought into the same form as matrix \( A \) in (11) by means of symmetric permutations (row and column interchanges).

In this talk, we first present eigenvalue bounds for \( P_{\text{convD}}^{-1} A \) and \( P_{\text{convT}}^{-1} A \) where

\[
\begin{align*}
P_{\text{convD}} &= \begin{bmatrix} A_{11} & 0 & 0 \\ 0 & A_{22} & B^T \\ 0 & B & 0 \end{bmatrix} \\
P_{\text{convT}} &= \begin{bmatrix} A_{11} & 0 & 0 \\ A_{21} & A_{22} & B^T \\ 0 & B & 0 \end{bmatrix}.
\end{align*}
\]

Field-of-values (FOV) type estimates for the above preconditioners have been already derived in [3]. For linear systems arising from finite element discretization of coupled Stokes-Darcy flow, we further establish upper and lower bounds for the eigenvalues of \( P_{T_1,\rho}^{-1} A \) where

\[
P_{T_1,\rho} := P_{T_1}(\rho) = \begin{bmatrix} A_{11} & 0 & 0 \\ 0 & A_{22} & 0 \\ 0 & B & -\rho M_p \end{bmatrix},
\]

here \( M_p \) is the mass matrix coming from the Stokes pressure; see [2, 3] for more details.

We also study the performance of following preconditioner:

\[
P_{r,\alpha} = \begin{bmatrix} A_{11} & A_{12} & 0 \\ 0 & A_{22} + rB^TQ^{-1}B & B^T \\ 0 & 0 & -\frac{1}{\alpha}Q \end{bmatrix}, \quad (12)
\]

for the augmented linear system \( \tilde{A}u = \tilde{b} \) in which

\[
\tilde{A} = \begin{bmatrix} A_{11} & A_{12} & 0 \\ A_{21} & A_{22} + rB^TQ^{-1}B & 0 \\ 0 & B & 0 \end{bmatrix},
\]

1Department of Mathematics, Vali-e-Asr University of Rafsanjan, PO Box 518, Rafsanjan, Iran
2Classe di Scienze, Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy
\( \hat{b} = (b_1; b_2 + rB^TQ^{-1}b_3; b_3) \), where the symmetric positive definite matrix \( Q \), the parameters \( r \geq 0 \) and \( \alpha > 0 \) are given. In particular, we either assume that \( r = \alpha \) or \( r = 0 \). Eigenvalue bounds for \( P_{0,\alpha}^{-1}A \) can be derived with a similar strategy used in [1, Theorem 2]. We comment that \( P_{0,\alpha} \) is the natural extension of a preconditioner proposed in [4, subsection 3.2]. Here, lower and upper bounds for the eigenvalues of preconditioned matrix \( P_{r,r}^{-1}A \) are obtained. Moreover, we establish FOV-type bounds for the preconditioned system associated with the preconditioner type \([12]\).

For 3D problems appearing in fluid mechanics, based on the results reported in [3], the preconditioners \( P_{\text{cond}} \) and \( P_{T_1,\rho} \) work better than the other proposed block preconditioners in the inexact implementation. In our implementation, we solved the subsystems required by the preconditioners \( P_{\text{cond}}, P_{T_1,\rho} \) and \( P_{r,r} \) by the PCG method (with early termination). We observed that unlike \( P_{\text{cond}} \), the outer iteration number for FGMRES in conjunction with \( P_{r,r} \) (and \( P_{T_1,\rho} \)) remains almost fixed as the grid is refined.

Numerical experiments will be reported for test problems from two different applications, the finite element discretization of coupled Stokes-Darcy flow and finite element modeling of liquid crystal directors, in order to compare the performance of inexact variants of these preconditioners.

References


Block Preconditioners for Incompressible Magnetohydrodynamics

Chen Greif[1]
Michael Wathen[2]

We consider preconditioning techniques for a mixed finite element discretization of an incompressible magnetohydrodynamics (MHD) problem. Upon discretization and linearization, a 4-by-4 nonsymmetric block-structured linear system needs to be (repeatedly) solved. One of the principal challenges is the presence of a skew-symmetric term that couples the fluid velocity with the magnetic field. We propose a preconditioner that exploits the block structure of the underlying linear system, utilizing and combining effective solvers for the mixed Maxwell and the Navier–Stokes subproblems. We perform a spectral analysis for an ideal version of the preconditioner, and develop and test a practical version of it. Large-scale numerical results in two and three dimensions validate the effectiveness of our approach.

[1] The University of British Columbia
[2] Rutherford Appleton Laboratory
When computing a very good preconditioner $P$ for a matrix $A$, we assume $P$ is a good approximation of $A^{-1}$; in other words, that $AP$ is a small perturbation of the identity matrix. Convergence rates for matrices of the forms $I + K$ and $I + E$ are well known (where $I$ is the identity, $\text{rank}(K) = p \ll n$, $\|E\|$ is small, and $\| \cdot \| = \| \cdot \|_2$). In this talk, we consider matrices of the form $I + K + E$. It has been shown that when the convergence tolerance is chosen to be larger than $\|E\|$, GMRES will converge in at most $p + 1$ iterations [1]. Rather than placing assumptions on the size of $\|E\|$, we consider how $\|E\|$ affects the eigenvalues of $I + K$. In particular, we consider matrices of the form $I + K + E$ within the context of the theoretical framework provided in [2], which analyzes the GMRES convergence of general perturbed matrices, $A + E$, by considering the pseudospectrum of $A$. Defining $A = I + K$, we examine the pseudospectrum of $I + K$, and in particular the sensitivity of the eigenvalues of $I + K$ to the introduction of $E$.

References


1Virginia Tech, Lehigh University
2Virginia Tech
We consider Anderson extrapolation \cite{Anderson1965} to accelerate the (stationary) Richardson iterative method for sparse linear systems. Using an Anderson mixing at periodic intervals, we assess how this benefits convergence to a prescribed accuracy. The method, named Alternating Anderson-Richardson \cite{Pratapa2019}, has appealing properties for high-performance computing, such as the potential to reduce communication and storage in comparison to more conventional linear solvers. We establish sufficient conditions for convergence, and we evaluate the performance of this technique in combination with various preconditioners through numerical examples. Furthermore, we propose an augmented version of this technique.

More details can be found in \cite{Pasini2019}.

References

\begin{itemize}
\end{itemize}
Preconditioning-based techniques for the convergence analysis of singularly perturbed convection-diffusion problems

Thái Anh Nhan\textsuperscript{1}
Relja Vulanović\textsuperscript{2}

We consider numerical methods for solving the linear singularly perturbed convection-diffusion problem,
\begin{equation}
Lu := -\epsilon u'' - b(x)u' + c(x)u = f(x), \; x \in (0,1), \; u(0) = u(1) = 0,
\end{equation}
where $\epsilon$ is a small positive perturbation parameter, $0 < \epsilon \ll 1$. In general, the solution $u$ has an exponential boundary layer near $x = 0$.

The problem (13) is discretized using finite-difference schemes on layer-adapted meshes. Generally speaking, such discretizations are not consistent uniformly in $\epsilon$, so $\epsilon$-uniform convergence cannot be proved by the classical approach based on $\epsilon$-uniform stability and $\epsilon$-uniform consistency. This is why previous proofs of convergence have introduced non-classical techniques (e.g., specially chosen barrier functions).

In this talk, we summarize our newly developed preconditioning-based approach—a suitable preconditioning of the discrete system is shown to yield a method that, uniformly in $\epsilon$, is both consistent and stable. Using this technique, $\epsilon$-uniform error bounds are obtained for the upwind \cite{4, 3} and hybrid higher-order finite-difference schemes \cite{2, 1}.

References


\textsuperscript{1}Department of Mathematics and Science, Holy Names University, 3500 Mountain Blvd., Oakland, CA 94619, USA
\textsuperscript{2}Department of Mathematical Sciences, Kent State University at Stark, 6000 Frank Ave NW, North Canton, OH 44720, USA
Orthogonalization schemes are critical for the convergence and the performance of nonsymmetric iterative solvers. (Typically Arnoldi-based Krylov methods, like for example GMRES.) In this talk we revisit the use of Householder reflections in the context of nonsymmetric Krylov solvers with one right-hand side (e.g. GMRES) or with more than one right-hand sides (e.g. Block GMRES).

In 1986, Saad and Schultz [1] introduces the GMRES algorithm. In 2005, Giraud, Langou, Rozložník, and van den Eshof [2] proved that CGS2 (Classical Gram-Schmidt with one re-orthogonalization) returned a $Q$ factor which is orthogonal up to machine precision. In 2006, C. C. Paige, M. Rozložník, and Z. Strakoš [3] proved that, despite severe loss of orthogonality in the $Q$ factor, MGS-GMRES (Modified Gram-Schmidt) is a backward stable algorithm. These two results show that CGS2 or MGS can safely be used in the context of GMRES with one right-hand side. In the block case (more than one right-hand sides), the stability of block GMRES with block-CGS2 or block-MGS is much less clearer.

In 1988, Walker [4] introduces the GMRES algorithm with Householder reflections (HH) as the orthogonalization scheme. Compared to Gram-Schmidt methods, Householder reflections have the advantage to be unconditionally stable as demonstrated by Wilkinson in 1965 [5]. In this talk, we revisit the HH-GMRES algorithm of Walker. Our main interest is in the block case. That being said, we believe that the single right-hand side case can also be well suited for HH-GMRES.

Firstly, we show the implementation of a one-synchronization GMRES Householder algorithm. By one-synchronization, we mean that the orthogonalization scheme only requires one global synchronization per step of the GMRES solver in a parallel distributed environment where vectors are distributed by rows. More synchronizations are needed due to the matrix-vector products. We only look at the orthogonalization scheme, (not the matrix-vector products,) and we claim that the orthogonalization scheme can be coded in such a way that it only requires one global synchronization. We do not look into the interaction (and possible communication savings) between the orthogonalization scheme and the matrix-vector products.

Secondly, we present an interface for orthogonalization scheme. This interface makes it easy to develop an iterative solver around an ad-hoc orthogonalization scheme. We believe most orthogonalization schemes can be encoded beneath this interface. The goal is to make the orthogonalization scheme independent from the iterative solver and to make many orthogonalization schemes available to any iterative solvers in a convenient way. We demonstrate this interface in the context of block methods and the Householder orthogonalization scheme. The interface is powerful and has implication beyond iterative solver by being relevant to dense
linear algebra libraries. We show how this interface enables to perform dense Householder QR factorization with many levels of blocking or with recursion. In 2000, Elmroth and Gustavson [6] demonstrated the benefits of using recursion in the context of dense Householder QR factorization for tall-and-skinny matrix. However for not-so-tall-and-skinny matrix, Householder QR with recursion has been deemed a challenge [6, 7] because of an unacceptable increase in the number of flops. Our new interface enables to use recursion while controlling the number of extra flops made by the algorithm and therefore is able to handle not-so-tall-and-skinny matrices.

**Thanks.** This work was sponsored by NSF award #1645514.

**References**


Low complexity matrix projections preserving actions on vectors

S. Cipolla 1
C. Di Fiore 2
P. Zellini 3

The projection onto algebras of matrices simultaneously diagonalized by unitary transforms $U$

$$\mathcal{L} := \text{sd } U = \{ U \text{Diag}(z) U^H : z \in \mathbb{C}^n \},$$

has been used profitably in the last thirty years as a core instrument in order to speed up, through preconditioning techniques, iterative methods for linear systems $A\mathbf{x} = \mathbf{b}$, where $A$ is a symmetric positive definite matrix. The main idea connected with these spaces could be traced in the key observation that, often, the matrices corresponding to linear systems arising from applications, exhibit some special structures and thus can be naturally approximated in low complexity spaces $\mathcal{L}$ of matrices of the form $\text{sd } U$. In particular $\mathcal{L}_A$, the projection in Frobenius of $A$ onto such $\mathcal{L} = \text{sd } U$, has revealed to produce approximations of the spectrum of $A$ good enough (see for example [1, 2, 3, 4]) to make $\mathcal{L}_A^{-1} A \approx I$ and thus to speed up preconditioned iterative solvers of $A\mathbf{x} = \mathbf{b}$ without increasing the time per step and the space complexity. This talk will show how, given a $n \times n$ symmetric matrix $A$, a matrix $V$ with $r$ orthonormal columns and an integer $m \geq 1$, $mr \leq n$, it is possible to devise a matrix algebra $\mathcal{L}$ such that one has $\mathcal{L}_A^j V = A^j V$ for $j = 0, \ldots, m - 1$. The algebra $\mathcal{L}$ is the space of all matrices that are diagonalized by a given orthogonal matrix $L$ which is the product of $mr$ Householder matrices, so that $\mathcal{L}$, for $mr \ll n$, is a low complexity matrix algebra. The fact that $\mathcal{L}_A^{-1} A^j V = V$, suggests to investigate the use of $\mathcal{L}_A$ as a new possible preconditioner for the Conjugate Gradient [5].

References


1University of Padua, Via Trieste 63, 35121 Padua (Italy) cipolla@math.unipd.it
2University of Rome “Tor Vergata”, Via Della Ricerca Scientifica 1, 00133 Rome (Italy)
3University of Rome “Tor Vergata”, Via Della Ricerca Scientifica 1, 00133 Rome (Italy)
Shift-without-invert and Shift-invert Techniques in Spectrum-partition for Accelerating Eigenvalue Calculations

Yunkai Zhou

We present a framework of spectrum partition based on ARPACK for accelerating large symmetric eigenvalue problems. Two methods within this framework are discussed: One utilizes shift-without-invert, which is less traditional but easier to implement; the other utilizes the more traditional shift-with-invert techniques. We discuss ways for automatic partition of unknown spectrum, adaptive choices of shifts, and ways to naturally reduce the cost in orthogonalization of eigenvectors.

\[^{1}\text{Southern Methodist University, yzhou@smu.edu}\]