



International Conference On Preconditioning Techniques For Scientific and Industrial Applications

June 10–12, 2024
Georgia Institute of Technology
Atlanta, USA

Sponsors:



Table of Contents

1. <u>Schedule</u>	2-3
2. <u>Directions to Conference Location</u>	4-5
3. <u>Program</u>	7-15
4. <u>Abstracts</u>	16-51
5. <u>Nearby Restaurants</u>	52
6. <u>Directions to Conference Dinner</u>	53

Schedule at a glance

Day 1: June 10

- 7:30 Registration and Breakfast
- 8:30 Opening Remarks
- 8:45 IP1
- 9:30 IP2
- 10:15 Coffee Break
- 10:45 CP1, CP2, CP3 (3 talks in 3 parallel sessions)
- 12:00 Lunch (attendees on own)
- 1:30 IP3
- 2:15 Intermission
- 2:25 MS1, MS2, MS3 (4 talks in 3 parallel sessions)
- 4:05 Coffee Break
- 4:35 CP4, CP5, CP6 (3 talks in 3 parallel sessions)
- 5:50 Adjourn for day 1

Day 2: June 11

- 7:30 Breakfast
- 8:45 IP4
- 9:15 IP5
- 10:15 Coffee Break and Group Photo
- 10:45 MS4, MS5, MS6 (3 talks in 3 parallel sessions)
- 12:00 Lunch (attendees on own)
- 1:30 IP6
- 2:15 Intermission

- 2:25 MS7, MS8, MS9 (4 talks in 3 parallel sessions)
- 4:05 Coffee Break
- 4:35 Panel Discussion
- 5:50 End of Panel Discussion
- 6:30 Dinner at South City Kitchen Midtown (1144 Crescent Ave NE, Atlanta)

Day 3: June 12

- 7:30 Breakfast
- 8:30 IP7 (15 minutes earlier than other days)
- 9:15 Intermission
- 9:25 MS10, MS11, MS12 (3 talks in 3 parallel sessions)
- 10:40 Coffee Break (20 minutes)
- 11:00 MS13, CP7 (3 talks in 2 parallel sessions)
- 12:15 Adjourn for day 3

Klaus Advanced Computing Building (KACB)

If you are at Spring St and 5th Street, the KACB building is a 10 minute walk. You can also take the free Gold Shuttle:

<http://bus.gatech.edu/routes/9/stops>

The above online/web map has “live shuttle locations” and that you should only board the Gold route buses.



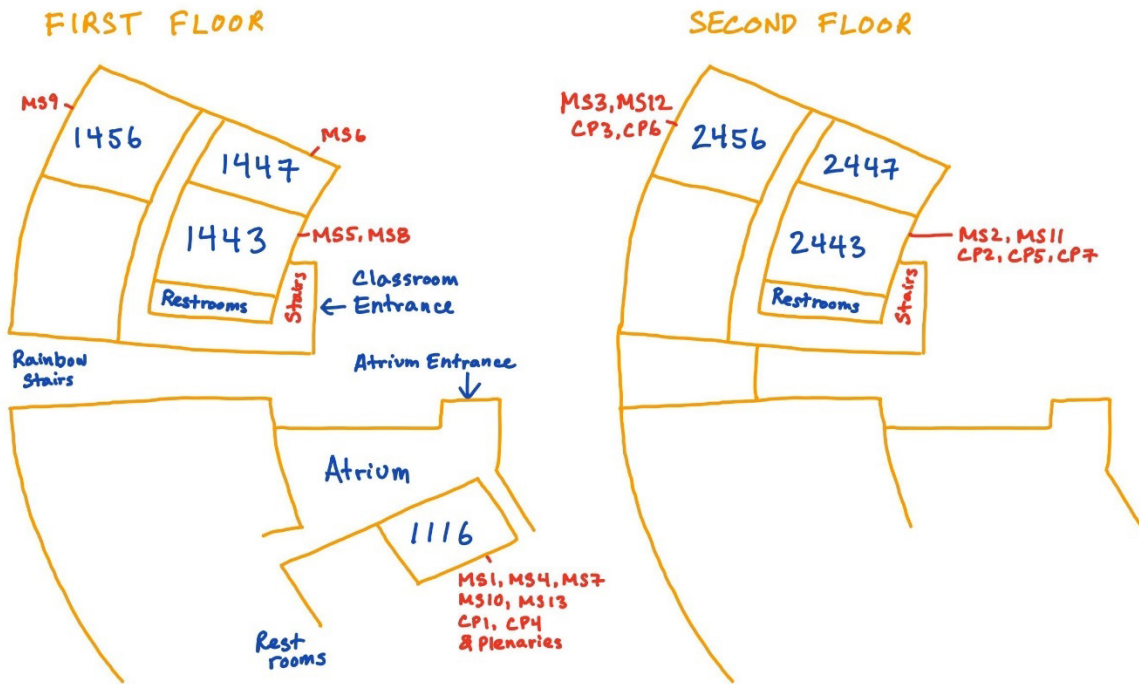
View of the KACB building from Ferst Dr. Proceed up the rainbow stairs.

If you do not want to take the stairs, you can take the path to the left to circle around the building.

The Research Wing entrance is on the right.



The entrance to the Classroom Wing is on the left.



To access rooms 14xx and 24xx, exit the Atrium and enter the Classroom wing building. For 24xx classrooms, take the stairs/elevator up to the second floor. The classrooms are on the right after passing the restrooms.

Precond24 Program

Schedule for Day 1: June 10, 2024

7:30--8:30	Registration and breakfast [Room: Atrium]	
8:30--8:45	Opening remarks [Edmond Chow, Yuanzhe Xi, Room: 1116]	
8:45--9:30	Invited Presentation - IP1 [Chair: Yousef Saad, Room: 1116] Machine Learning and Multilevel Methods: The Integration Rui Peng Li Lawrence Livermore National Laboratory, USA	
9:30--10:15	Invited Presentation - IP2 [Chair: Yousef Saad, Room: 1116] Matrix-free solvers for high-order, high-performance finite elements Will Pazner Portland State University, USA	
10:15--10:45	Coffee Break [Room: Atrium]	
Parallel Contributed Presentation Sections (10:45-12:00)		
CP1 Preconditioners for physical problems Chair: <i>Daniel Osei-Kuffuor</i> Room: 1116	CP2 Machine Learning enhanced preconditioners Chair: <i>Yuanzhe Xi</i> Room: 2443	CP3 Preconditioners for structured matrices Chair: <i>Edmond Chow</i> Room: 2456

<p>A New Preconditioned Nonlinear Conjugate Gradient Method in Real Arithmetic for Computing the Ground States of Rotational Bose-Einstein Condensate 10:45--11:10</p> <p>Fei Xue Clemson University</p>	<p>Spectral-Refiner: Fine-tuning for accurate spatiotemporal operator learning in turbulent flows 10:45--11:10</p> <p>Francesco Brarda Emory University</p>	<p>Near-optimal hierarchical matrix approximation from matrix-vector products 10:45--11:10</p> <p>Tyler Chen New York University</p>
<p>Preconditioning finite difference Hamiltonian matrices from density functional theory 11:10--11:35</p> <p>Shikhar Shah Georgia Tech</p>	<p>Augmenting linear solvers in fusion codes with neural operators 11:10--11:35</p> <p>Yang Liu Lawrence Berkeley National Lab</p>	<p>Multigrid method for hierarchical rank structured matrices 11:10--11:35</p> <p>Daria Sushnikova King Abdullah University of Science and Technology</p>
<p>Achieving Additive Separability in the Internally Contracted Multireference Unitary Coupled-Cluster Method Through the Inexact Newton Approach 11:35--12:00</p> <p>Shuhang Li Emory University</p>	<p>Adaptive factorized Nystrom preconditioner for kernel matrices 11:35--12:00</p> <p>Hua Huang Georgia Tech</p>	<p>Efficient SAA Methods for Hyperparameter Estimation in Bayesian Inverse Problems 11:35--12:00</p> <p>Malena Sabaté Landman Emory University</p>
Lunchtime: 12-1:30 (attendees on own)		
<p>1:30--2:15</p>	<p>Invited Presentation - IP3 [Chair: Wil Schilders, Room: 1116]</p> <p>A fine-grained fully iterative ILU preconditioner for unsteady density variable Navier-Stokes equations</p> <p>Monica Dessolet CERN SFT-EP, Switzerland</p>	
<p>2:15--2:25</p>	Intermission	
Parallel Minisymposium Sections (2:25-4:05)		

MS1 Parallel and Machine Learning Preconditioning Methods for Large Linear Systems Chairs: <i>Kees Vuik and Wil Schilders</i> Room: 1116	MS2 Preconditioners for High-Frequency Helmholtz Problems Chairs: <i>Kars Knook, Robert Kirby, and Drew Anderson</i> Room: 2443	MS3 Recent advances in multigrid preconditioning Chairs: <i>Victor Magri and Nicola Castelletto and Daniel Osei-Kuffuor</i> Room: 2456
Prospective on Latest Advances of Scalable Hybrid Monte Carlo Methods for Linear Algebra 2:25--2:50 Vassil Alexandrov Hartree Centre	Towards scalable preconditioners for indefinite systems arising in electromagnetic simulations 2:25--2:50 Vandana Dwarka TU Delft	An interior-point multigrid-based approach for scalable computational contact mechanics 2:25--2:50 Tucker Hartland Lawrence Livermore National Laboratory
Matrix-Free Parallel Scalable Multilevel Deflation Preconditioning for the Helmholtz Equation 2:50--3:15 Jinqiang Chen TU Delft	Using Spectral Coarse Spaces of the H-Geneo Type for Efficient Solutions of the Helmholtz Equation 2:50--3:15 Victorita Dolean TU Eindhoven	Robust physics-based preconditioners for multi-physics problems 2:50--3:15 Xiaozhe Hu Tufts University
Scalable distributed preconditioners in Ginkgo 3:15--3:40 Pratik Nayak Karlsruhe Institute of Technology	Acceleration of non-local exchange in generalized optimized Schwarz methods 3:15--3:40 Xavier Claeys Sorbonne Université	A multigrid reduction framework for multi-physics applications 3:15--3:40 Victor Magri Lawrence Livermore National Laboratory
DeepONet-based Preconditioning for Krylov Methods 3:40--4:05 Alena Kopanicakova Università della Svizzera italiana	Some convergence results for RAS-Imp and RAS-PML for the Helmholtz equation 3:40--4:05 Shihua Gong The Chinese University of Hong Kong, Shenzhen, and SICIAM, SRIBD	Mixed precision algorithm development in hypre 3:40--4:05 Ulrike Yang Lawrence Livermore National Laboratory

4:05--4:35	Coffee Break [Room: Atrium]	
Parallel Contributed Presentation Sections (4:35-5:50)		
CP4 Multigrid preconditioners Chair: <i>Ulrike Yang</i> Room: 1116	CP5 Domain decomposition preconditioners Chair: <i>Nicole Spillane</i> Room: 2443	CP6 Advanced Preconditioning Techniques Chair: <i>Erik Boman</i> Room: 2456
<u>Multigrid for block Toeplitz systems arising from PDEs and systems thereof</u> 4:35--5:00 Matthias Bolten Bergische Universität Wuppertal	<u>Brain edema simulation by using domain decomposition methods</u> 4:35--5:00 Talal Alshehri Morgan State University	<u>Preconditioning Techniques for Multiterm Generalized Sylvester Equations</u> 4:35--5:00 Yannis Voet École Polytechnique Fédérale de Lausanne
<u>Symbol-Based Analysis of (Two Related) Multigrid Methods for Electromagnetic Scattering Problems</u> 5:00--5:25 René Spoerer Bergische Universität Wuppertal	<u>Robust Domain Decomposition Methods for High-contrast Multiscale Problems on Irregular Domains</u> 5:00--5:25 Juan Calvo University of Costa Rica	<u>Preconditioning for Topological Constraint Problem</u> 5:00--5:25 Mingdong He University of Oxford
<u>Practical Advances in High-Order Stokes Solvers: Robust Multigrid Preconditioners with AMG Integration</u> 5:25--5:50 Alexey Voronin University of Illinois Urbana Champaign	<u>Preconditioned IDR Solution Methods in Scientific and Industrial Applications</u> 5:25--5:50 Alex Fedoseyev Ultra Quantum Inc.	<u>Data-Driven Solver and Preconditioner Selection for Sparse Linear Matrices</u> 5:25--5:50 Hayden Liu Weng Technical University of Munich

Schedule for Day 2: June 11, 2024

7:30--8:45	Breakfast [Room: Atrium]	
8:45--9:30	Invited Presentation - IP4 [Chair: Esmond Ng, Room: 1116] The importance of coarse levels for domain decomposition methods Alexander Heinlein Delft University of Technology, Netherlands	
9:30--10:15	Invited Presentation - IP5 [Chair: Esmond Ng, Room: 1116] Scaled spectral preconditioners for sequence of linear systems with an application to data assimilation Selime Gürol CERFACS, France	
10:15--10:45	Coffee Break and Group Photo [Room: Atrium]	
Parallel Minisymposium Sections (10:45-12:00)		
MS4 Analog and mixed precision preconditioning Chairs: <i>Erik Boman, Mark S. Squillante, Vassilis Kalantzis</i> Room: 1116	MS5 Recent Advances in Saddle-Point and Double Saddle-Point Systems Chair: <i>Chen Greif</i> Room: 1443	MS6 Nonlinear Preconditioning Techniques and Applications I Chairs: <i>Xiao-Chuan Cai and Alexander Heinlein</i> Room: 1447
Solvers and Preconditioners for Analog Architectures 10:45--11:10 Erik Boman Sandia National Laboratories	Spectral Properties of Double Saddle-Point Systems 10:45--11:10 Chen Greif The University of British Columbia	Some preconditioned inexact Newton methods with learning capabilities 10:45--11:10 Xiao-Chuan Cai University of Macau

<p>Solving Sparse Linear Systems via Flexible GMRES with In-Memory Analog Preconditioning</p> <p>11:10--11:35</p> <p>Chai Wah Wu</p> <p>IBM Research</p>	<p>Block Triangular Preconditioners for Double Saddle-Point Problems Arising in Mixed Hybrid Coupled Poromechanics</p> <p>11:10--11:35</p> <p>Massimiliano Ferronato</p> <p>University of Padova</p>	<p>Domain decomposition preconditioners and multi-scale approaches to solve stationary and time-dependent nonlinear equations</p> <p>11:10--11:35</p> <p>Victorita Dolean</p> <p>TU Eindhoven</p>
<p>Half precision wave simulation</p> <p>11:35--12:00</p> <p>Longfei Gao</p> <p>Argonne National Laboratory</p>	<p>An Augmented Lagrangian Preconditioner for the Control of the Navier--Stokes Equations</p> <p>11:35--12:00</p> <p>Santolo Leveque</p> <p>Scuola Normale Superiore di Pisa</p>	<p>Nonlinear Preconditioning for Implicit Solution of Discretized PDEs</p> <p>11:35--12:00</p> <p>David Keyes</p> <p>King Abdullah University of Science and Technology</p>
<p>Lunchtime: 12-1:30 (attendees on own)</p>		
<p>1:30--2:15</p>	<p>Invited Presentation - IP6 [Chair: Andy Wathen, Room: 1116]</p> <p>Leveraging multipreconditioning for the efficient solution of High-Frequency Helmholtz problems</p> <p>Tyrone Rees</p> <p>STFC, UK</p>	
<p>2:15--2:25</p>	<p>Intermission</p>	
<p>Parallel Minisymposium Sections (2:25-4:05)</p>		
<p>MS7 Preconditioned Linear Algebraic Techniques for Solving Inverse Problems</p> <p>Chairs: <i>Lucas Onisk and James Nagy</i></p> <p>Room: 1116</p>	<p>MS8 Algebraic and Geometric Domain Decomposition Preconditioners for Complex Problems</p> <p>Chairs: <i>Victorita Dolean and Nicole Spillane</i></p> <p>Room: 1443</p>	<p>MS9 Preconditioning and Machine Learning I</p> <p>Chairs: <i>Ye Qiang and Jianlin Xia</i></p> <p>Room: 1456</p>

Effective Approximate Preconditioners for Linear Inverse Problems 2:25--2:50 Lucas Onisk Emory University	Substructuring the Hiptmair-Xu Preconditioner 2:25--2:50 Xavier Claeys Sorbonne Université	Batch Normalization Preconditioning for Neural Network Training 2:25--2:50 Qiang Ye University of Kentucky
A New Deflation Space for Preconditioned GMRES 2:50--3:15 Daniel Szyld Temple University	Overlapping Schwarz Preconditioner with Geneo Coarse Space for Nonlocal Equations 2:50--3:15 Pierre Marchand INRIA Paris Saclay	Batch Normalization Preconditioning for Convolutional Neural Networks 2:50--3:15 Susanna Lange University of Chicago
Preconditioning Linear Inverse Problems Using Randomization and Subspace Projection 3:15--3:40 Eric de Sturler Virginia Tech	An Algebraic Domain Decomposition Preconditioner 3:15--3:40 Nicole Spillane CNRS, Ecole Polytechnique	A Gromov--Wasserstein Geometric Objective for Graph Coarsening and Potentials for Preconditioning 3:15--3:40 Jie Chen MIT-IBM Watson AI Lab
Randomized Approaches for Optimal Experiment Design 3:40--4:05 Srinivas Eswar Argonne National Laboratory	Development of preconditioning techniques for integrated energy systems 3:40--4:05 Buu-Van Nguyen Delft University of Technology	A structure-guided Gauss-Newton method for shallow ReLU neural network 3:40--4:05 Tong Ding Purdue University
4:05--4:35	Coffee Break [Room: Atrium]	
4:35--5:50	Panel Discussion [Room 1116]: Machine Learning and Numerical Linear Algebra Chair: Edmond Chow Panelists: Jie Chen, Victorita Dolean, Lars Ruthotto and Qiang Ye	
6:30--	Banquet at South City Kitchen Midtown	

Schedule for Day 3: June 12, 2024

7:30--8:30	Breakfast [Room: Atrium]		
8:30--9:15	Invited Presentation - IP7 [Chair: Andy Wathen, Room: 1116] TorchBraid: High-Performance Layer-Parallel Training of Deep Neural Networks with MPI and GPU Acceleration Jacob B. Schroder University of New Mexico, USA		
9:15--9:25	Intermission		
Parallel Minisymposium Sections (9:25-10:40)			
MS10 Preconditioning and Machine Learning II Chairs: <i>Qiang Ye and Jianlin Xia</i> Room: 1116	MS11 Nonlinear Preconditioning Techniques and Applications II Chairs: <i>Xiao-Chuan Cai and Alexander Heinlein</i> Room: 2443	MS12 Preconditioning Techniques for Gaussian Processes Chairs: <i>Tianshi Xu and Mikhail Lepilov</i> Room: 2456	
Equivariant Generative Models for Molecular Modeling 9:25--9:50 Bao Wang University of Utah	Exploring nonlinear preconditioning strategies for solving phase-field fracture problems 9:25--9:50 Hardik Kothari Università della Svizzera italiana	On Gaussian Kernel Matrices: Spectral Properties and Efficient Approximations 9:25--9:50 Difeng Cai Southern Methodist University	
Generating Polynomial Method for Non-symmetric Tensor Decomposition 9:50--10:15 Zequn Zheng Louisiana State University	Accelerating training of physics-informed neural networks using decomposition strategies 9:50--10:15 Alena Kopanicakova Università della Svizzera italiana	Spectral Shape Estimation of Kernel Matrices 9:50--10:15 Mikhail Lepilov Emory University	

Fast solvers for neural network least-squares approximations 10:15--10:40 Jianlin Xia Purdue University	Adaptive optimised Schwarz methods 10:15--10:40 Conor McCoid Université Laval	Efficient Preconditioned Unbiased Estimators in Gaussian Processes 10:15--10:40 Tianshi Xu Emory University
10:40--11:00	Coffee Break [Room: Atrium]	
Parallel Sections (11:00-12:15)		
MS13 Recent Progress on Learning to Precondition with Graph Neural Networks Chair: <i>Jie Chen</i> Room: 1116	CP7: Advances in Multigrid Preconditioners Chair: <i>Chen Greif</i> Room: 2443	
Graph neural network based preconditioner for Krylov subspace methods 11:00--11:25 Paul Häusner Uppsala University	Nesting Approximate Inverses for Improved Preconditioning and Algebraic Multigrid Smoothing 11:00--11:25 Andrea Franceschini University of Padova	
Graph Neural Networks for Selection of Preconditioners and Krylov Solvers 11:25--11:50 Ziyuan Tang University of Minnesota	LFA-tuned matrix-free multigrid for the elastic Helmholtz equation 11:25--11:50 Rachel Yovel Ben-Gurion University of the Negev	
Approximating the Inverse of a Sparse Linear Operator with Graph Neural Networks 11:50--12:15 Jie Chen MIT-IBM Watson AI Lab, IBM Research	Bi-parametric Operator Preconditioning 11:50--12:15 Carlos Jerez-Hanckes University of Bath, Universidad Adolfo Ibáñez	

Abstracts

Invited Talks:

IP 1: Machine Learning and Multilevel Methods: The Integration

Rui Peng Li

Lawrence Livermore National Lab, USA

Abstract:

In this presentation, we will share our ongoing exploration of the synergy between machine learning and multilevel methods. Our research has progressed in two directions, each leveraging one approach to enhance the other. Multilevel methods are inherently complex, requiring complementary operators to ensure overall efficiency. Designing these algorithms typically demands careful customization for specific application problems. Our objective is to leverage machine learning techniques to automatically discover more efficient and robust multigrid algorithms. Conversely, given that neural network training is challenging and computationally expensive, we aim to utilize multilevel methods to enhance training efficiency and stability. First, we will discuss employing machine learning-based methods to construct robust operators for multigrid solvers, including neural-network-based smoothers and machine learning approaches for non-Galerkin coarse-grid operators. Second, we will explore enhancing neural network training through the integration of nonlinear multigrid methods. This involves multiple levels of NNs with decreasing complexities collaborating to train the largest NN at the finest level, with parameters being transferred, optimized, and corrected at each level. Additionally, we will present numerical results from scientific computing to substantiate our findings and demonstrate the practical applications of our research.

IP 2: Matrix-free solvers for high-order, high-performance finite elements

Will Pazner

Portland State University, USA

Abstract:

High-order discretizations result in highly accurate, predictive simulations, and can achieve high performance on emerging computing architectures, such as GPU-based exascale supercomputers. However, efficiently solving the resulting systems can be challenging; these systems are denser and more ill-conditioned than those of low-order discretizations. Assembling and storing the system matrix is often prohibitively expensive, and the convergence of traditional solver techniques such as algebraic multigrid may not be satisfactory when applied to these problems. In this talk, I will discuss the development of matrix-free, high-performance, GPU-accelerated preconditioners for a broad class of high-order finite element problems. The core idea of these preconditioners is the construction of a spectrally equivalent low-order discretization posed on a refined mesh; this is a classical idea, first proposed by Orszag in 1980. Through the introduction of a polynomial basis using interpolation and histopolation operators, this approach can be extended to high-order problems in all spaces of the finite element de Rham complex. Properties of this basis can be exploited to construct high-performance saddle-point solvers for mixed finite element problems in $H(\text{div})$, including Darcy and grad-div problems. Robust preconditioners for discontinuous Galerkin discretizations are developed using similar techniques. These solvers deliver uniform convergence with respect to problem size and order of the method. The efficient implementation of these methods on GPU-based architectures will be discussed.

IP3: A fine-grained fully iterative ILU preconditioner for unsteady density variable Navier-Stokes equations

Monica Dessoie

CERN SFT-EP, Switzerland

Abstract:

Solving a sequence of slowly varying linear systems sharing the same sparsity pattern is a frequently encountered problem in many applications, the most notable being time-dependent PDEs. A fine-grained fully iterative ILU preconditioning strategy is here presented to cope with such problem. The algorithm discussed includes an iterative updating strategy, as well as an iterative method for solving the triangular systems that result from the ILU preconditioner. We analyse the performance of the proposed method in terms of robustness, number of iterations for convergence and time-to-solution, focusing on massively parallel accelerators, such as GPUs. In particular, we address the solution of incompressible flows with variable density, showing results for simulations of mixtures of immiscible liquids, i.e. the well-known Rayleigh–Taylor instability. We largely investigate the interplay between Reynolds and Atwood numbers, two adimensional quantities describing flow turbulence and fluid density ratio, respectively. We show how this fully iterative approach turns out to be robust and efficient for many configurations of this problem.

IP4. The importance of coarse levels for domain decomposition methods

Alexander Heinlein

Abstract:

Domain decomposition methods (DDMs) solve boundary value problems by decomposing them into smaller subproblems defined on an overlapping or non-overlapping decomposition of the computational domain. Their divide-and-conquer approach makes them well-suited for parallel computing. However, achieving robust convergence for challenging problems and scalability to large numbers of subdomains generally requires (global) information transport. This can be achieved by incorporating a well-designed coarse level, transforming DDMs from one- to multi-level algorithms. This talk highlights the importance of coarse levels in domain decomposition methods. First, the algorithmic framework of extension-based coarse spaces will be discussed. They provide robustness and scalability to Schwarz preconditioners for a wide range of challenging problems exhibiting, for instance, strong heterogeneities [2], multiple coupled physics [5], and/or strong nonlinearities [4]. Numerical results using the FROSch (Fast and Robust Overlapping Schwarz) package [3], which is part of the Trilinos library, demonstrate the effectiveness and efficiency of these Schwarz preconditioners. The second part of the talk will explore the application of DDMs to neural networks (NNs), demonstrating improvements in terms of accuracy, computation time, and/or memory efficiency. Similar to classical domain decomposition methods, coarse levels, here in the form of small global NNs, ensure global information transport, enabling scalability. This talk will cover the application of DDMs in solving partial differential equations using physics-informed NNs (PINNs) [1] and in image segmentation using convolutional NNs (CNNs) [6].

References

- [1] Victorita Dolean, Alexander Heinlein, Siddhartha Mishra, and Ben Moseley. Multilevel domain decomposition-based architectures for physics-informed neural networks. arXiv preprint arXiv:2306.05486, 2023.
- [2] Alexander Heinlein, Axel Klawonn, Jascha Knepper, Oliver Rheinbach, and Olof B. Widlund. Adaptive GDSW coarse spaces of reduced dimension for overlapping Schwarz methods. *SIAM Journal on Scientific Computing*, 44(3):A1176–A1204, 2022.
- [3] Alexander Heinlein, Axel Klawonn, Sivasankaran Rajamanickam, and Oliver Rheinbach. FROSch: A Fast and Robust Overlapping Schwarz domain decomposition preconditioner based on Xpetra in Trilinos. Springer, 2020.
- [4] Alexander Heinlein and Martin Lanser. Additive and hybrid nonlinear two-level Schwarz methods and energy minimizing coarse spaces for unstructured grids. *SIAM Journal on Scientific Computing*, 42(4):A2461–A2488, 2020.
- [5] Alexander Heinlein, Mauro Perego, and Sivasankaran Rajamanickam. FROSch preconditioners for land ice simulations of Greenland and Antarctica. *SIAM Journal on Scientific Computing*, 44(2):B339–B367, 2022.
- [6] Corne Verburg, Alexander Heinlein, and Eric. C. Cyr. A domain decomposition-based CNN for high-resolution image segmentation. in preparation.

IP5. Scaled spectral preconditioners for sequence of linear systems with an application to data assimilation

Selime Gürol

CERFACS, France

Abstract:

Computational science and engineering problems often require the solution of a sequence of symmetric linear systems. This situation arises, for example, in iterative solutions to nonlinear least squares problems, or in uncertainty quantification applications. In this study, we focus on the preconditioned Conjugate Gradient (PCG) method to solve each system. Typically, a first level preconditioner, denoted as F_1 , is used for the initial linear system, $A_1x_1 = b_1$. The choice of the first-level preconditioner depends on the specific problem and may consider the physical properties of the problem and/or the algebraic structure of the matrix A_1 . To further accelerate the rate of convergence of the PCG for subsequent linear systems $A_{j+1}x_{j+1} = b_{j+1}$, it is common to use such a low-rank update to the most recent preconditioner, F_j , leveraging information obtained from solving the previous linear system, i.e. $A_jx_j = b_j$. A particularly common choice to construct a low-rank update is to use the (approximate) eigenspectrum of the matrix A_j . The underlying idea is to capture the remaining eigenvalues after the application of the first-level preconditioning and then cluster them to a positive value, typically around 1. In this study, the emphasis will be on the scaled spectral preconditioner, which is defined by a scaling parameter determining the position of the cluster. We will present various strategies for the choice of this scaling parameter, as it plays a significant role in the convergence behavior of the PCG method. As certain applications exhibit computational constraints requiring the use of truncated CG, such as numerical weather forecast, we will focus also on the early convergence properties of the PCG to design more efficient preconditioner. Our theoretical results are validated with numerical experiments based on reference atmospheric models, such as Lorenz-96 or the quasi-geostrophic model within the Object Oriented System (OOPS) developed by Météo-France and ECMWF.

IP6. Leveraging multipreconditioning for the efficient solution of High-Frequency Helmholtz problems

Tyrone Rees

STFC Computational Mathematics Group

Abstract:

The design of preconditioners that can harness the power of modern, massively parallel computing architectures presents a significant challenge. One promising approach to enhancing parallelism in iterative linear solves is multipreconditioning, a technique that enables the simultaneous application of multiple preconditioners. This presentation will delve into the intricacies of multipreconditioning, with a special emphasis on its application to high-frequency Helmholtz problems. The finite element method discretization of Helmholtz problems results in complex, non-Hermitian sparse linear systems that are challenging to solve numerically. Our

work draws inspiration from domain decomposition strategies, particularly sweeping methods. These methods have garnered significant interest due to their ability to achieve nearly-linear asymptotic complexity, making them particularly suitable for high-frequency problems. We explore the use of straightforward sweeping techniques applied in various directions, which can be integrated in parallel into a multipreconditioned GMRES strategy. Each sweep involves solving smaller Helmholtz problems, each of which do not require highly accurate solutions. This allows for the potential of a matrix-free approach, as we can recursively apply the same strategy or any other effective Helmholtz solver. Numerical results will be presented to demonstrate the efficacy of our comprehensive solver strategy.

IP7. TorchBraid: High-Performance Layer-Parallel Training of Deep Neural Networks with MPI and GPU Acceleration

Jacob B. Schroder

University of New Mexico, USA

Abstract:

Deep neural networks (DNNs) exhibit excellent performance for many machine learning tasks, e.g., image classification, natural language processing, and game playing. However, training DNNs remains challenging and computationally expensive, with much room for improvement, both in terms of new sources of parallelism and algorithmic speedup. One of the key bottlenecks is the serialization inherent in forward and backward propagation, which limits strong scaling in the limit. Recently, the parallel-in-time method, multigrid-reduction-in-time (MGRIT), has been applied to some DNNs to overcome this bottleneck by providing new parallelism in the layer dimension (layer-parallelism). This new parallelism is made possible by a connection between the layer-dimension and a hypothetical time-dimension. In this talk, we introduce layer-parallelism with MGRIT and then discuss TorchBraid, which is a high-performance implementation of this approach that supports MPI-based parallelism in combination with GPU acceleration. To achieve this, TorchBraid integrates the PyTorch neural network framework with the XBraid time-parallel library. We present results for Torchbraid with and without GPU acceleration, considering Tiny ImageNet and MNIST, as well as recurrent neural networks and transformers for language processing. We also present new results showing the computational advantage of combining layer-parallelism with data-parallelism and how to adapt standard deep learning techniques, like batch-normalization, to the layer-parallel setting. Lastly, we discuss TorchBraid's approach for overcoming the algorithmic challenges inherent in combining automatic differentiation with layer-parallel in a distributed MPI setting. Overall, TorchBraid enables fast training of DNNs, both in a strong and weak scaling context.

Contributed Presentation Sections

CP 1. Preconditioners for physical problems

Title: A New Preconditioned Nonlinear Conjugate Gradient Method in Real Arithmetic for Computing the Ground States of Rotational Bose-Einstein Condensate

Fei Xue

Clemson University

Abstract:

We propose a new nonlinear preconditioned conjugate gradient (PCG) method in real arithmetic for computing the ground states of rotational Bose-Einstein condensate, modeled by the Gross-Pitaevskii equation. We show that the special structure of the energy functional $E(\phi)$ and its gradient with respect to ϕ can be fully exploited in real arithmetic. We propose a simple approach for fast evaluation of the energy functional, which enables exact line search. We derive the discrete Hessian operator of the energy functional and propose a shifted Hessian preconditioner for PCG. With our ideal Hessian preconditioner, PCG is expected to exhibit mesh size-independent asymptomatic convergence behavior. Numerical experiments in 2D and 3D domains show the efficiency of fast energy evaluation, the robustness of exact line search, and the improved convergence of PCG with our new preconditioner in iteration counts and runtime, notably for more challenging rotational BEC problems with high nonlinearity and rotational speed.

Title: Preconditioning finite difference Hamiltonian matrices from density functional theory

Shikhar Shah

Georgia Tech

Abstract:

We consider solving many block linear systems of the form $(A + z_i I)X = B_i$, where A is a fixed Hermitian matrix and each z_i is a complex constant. The matrix A is the sum of a high-order finite difference approximation to the Laplacian and a low-rank matrix. We precondition each block linear system using an efficient Poisson solve based on the Kronecker product formulation of the discrete Laplacian matrix. The effectiveness of this preconditioner is a function of both the block size used in the linear solve as well as the constant z_i . We numerically investigate this effect and determine the circumstances in which this preconditioner yields faster solution times compared to the unpreconditioned systems.

Title: Achieving Additive Separability in the Internally Contracted Multireference Unitary Coupled-Cluster Method Through the Inexact Newton Approach

Shuhang Li

Emory University

Abstract:

In this work, we present our advancements in the internally contracted multireference unitary coupled-cluster method (ic-MRUCC) and address the challenge of its applicability to molecular stretched geometries. The primary issue with the current ic-MRUCC implementation is its failure to achieve additive separability of energy, a critical property for accurately describing dissociated systems. To overcome this, we adopt a projective approach that leverages generalized normal ordering (GNO). This technique effectively circumvents the disconnect diagrams responsible for the breakdown of additive separability. Furthermore, we employ an orthogonalization procedure to deal with the linearly dependent excitation configuration basis, coupled with an inexact Newton method to solve the resulting nonlinear equations. Our computational results demonstrate that our ic-MRUCC implementation successfully restores additive separability, making it a robust tool for potential energy surface construction.

CP 2. Machine Learning enhanced preconditioners

Title: Spectral-Refiner: Fine-tuning for accurate spatiotemporal operator learning in turbulent flows

Francesco Brarda

Emory University

Abstract:

In this talk, we propose a new Spatiotemporal Fourier Neural Operator (SFNO) that learns maps between Bochner spaces. This new paradigm leverages wisdom from traditional numerical PDE theory and techniques to refine the pipeline of commonly adopted end-to-end neural operator training and evaluations. Specifically, in the learning problems for the turbulent flow modeling by the Navier-Stokes Equations (NSE), the proposed architecture initiates the training with a few epochs for SFNO, concluding with the freezing of most model parameters.

Then, the last linear spectral convolution layer is fine-tuned without the frequency truncation. The optimization uses a negative Sobolev norm for the first time as the loss in operator learning, defined through a reliable functional-type a posteriori error estimator whose evaluation is almost exact thanks to the Parseval identity. This design allows the neural operators to effectively tackle low-frequency errors while the relief of the de-aliasing filter addresses high-frequency errors. Numerical experiments on commonly used benchmarks for the 2D NSE demonstrate

significant improvements in both computational efficiency and accuracy, compared to end-to-end evaluation and traditional numerical PDE solvers.

Title: Augmenting linear solvers in fusion codes with neural operators

Yang Liu

Lawrence Berkeley National Lab

Abstract:

Realistic fusion simulation codes, typically fluid model or Gyrokinetic particle-in-cell method-based, are characterized by expensive execution, multiple variables, and rich nonlinear dynamics. This work develops Fourier neural operator (FNO)-based cheap surrogates for these simulation codes by combining the conventional FNO architecture with the semi-discretized governing equations of a given fusion code. This improved architecture, called fusion-FNO, is capable of significantly reducing the parameter counts while maintaining similar prediction accuracy. We demonstrate its efficiency and capability using Department of Energy's fusion codes NIMROD and GTC. We further demonstrate how to use fusion-FNO to augment linear solvers in fusion codes using a simple immersed boundary projection method-based code.

Title: Adaptive factorized Nystrom preconditioner for kernel matrices

Huang Hua

Georgia Tech

Abstract:

In this presentation, we will present robust preconditioning strategies for the iterative solution of systems involving kernel matrices. The characteristics of a kernel matrix, including its spectrum, are heavily influenced by the parameters of the kernel function, like the length scale. This dependency poses a challenge in designing a preconditioner that is effective across various parameter settings for a (regularized) kernel matrix. We will delve into the Nystrom approximation, a technique that proves highly effective for kernel matrices of low rank. For matrices of moderate rank, we propose an enhancement to the Nystrom method. The improved preconditioner, featuring a block-factorized structure, shows great efficiency even with kernel matrices that have large numerical ranks. Key aspects we will cover include estimating the kernel matrix's rank and selecting landmark points for the Nystrom approximation.

CP 3. Preconditioners for structured matrices

Title: Near-optimal hierarchical matrix approximation from matrix-vector products

Tyler Chen

New York University

Abstract:

A number of algorithms for recovering a hierarchical off-diagonal low-rank (HODLR) matrix A , accessed only from matrix-vector products, have been developed. How do these algorithms work for the more general problem of finding a HODLR approximation to an arbitrary matrix? We show certain variants of so-called "peeling algorithms" can provably obtain near-optimal approximations. We also provide numerical evidence that others could be exponentially unstable.

Title: Multigrid method for hierarchical rank structured matrices

Daria Sushnikova

King Abdullah University of Science and Technology

Abstract:

In the talk, a new fast solver designed for large systems with hierarchical block low-rank matrices is introduced. The algorithm combines H2 matrix approximations with the Multigrid method, creating a synthesis of the two: the H2-MG algorithm. This innovative combination combines the time and memory efficiency of the H2 matrix along with the fast convergence of Multigrid. The talk will explain the mechanics and theoretical foundation of the H2-MG algorithm, demonstrate its linear complexity, and highlight its effectiveness through several kernel matrices examples. While the current range of H2 solvers includes various effective iterative and direct methods, it notably lacks one that employs the Multigrid approach. The introduction of the H2-MG algorithm marks a significant addition to the business of H2 matrix solvers, offering a new direction for progress in fields dealing with large, dense, and ill-conditioned matrices.

Title: Efficient SAA Methods for Hyperparameter Estimation in Bayesian Inverse Problems

Malena Sabaté Landman

Emory University

Abstract:

In Bayesian inverse problems, there are several hyperparameters that define the prior and the noise model and must be estimated from the data. For linear inverse problems with additive Gaussian noise and Gaussian priors defined using Matern covariance models, we estimate the hyperparameters using the maximum a posteriori estimate of the marginalized posterior distribution. However, this is a computationally intensive task since it involves computing log determinants. To address this challenge, we consider a stochastic average approximation (SAA) of the objective function and use preconditioned Lanczos methods to efficiently approximate the objective function and the gradient. We demonstrate the performance of our

approach on synthetic and real data inverse problems from tomography and atmospheric transport.

CP 4. Multigrid preconditioners

Title: Multigrid for block Toeplitz systems arising from PDEs and systems thereof

Matthias Bolten

Bergische Universität Wuppertal

Abstract:

The discretization of PDEs in large sparse linear systems, in the case of structured meshes, suitable boundary conditions and constant coefficients the resulting matrices are Toeplitz matrices. Multigrid methods are known to be optimal solvers for many of these systems, yet their convergence has been mostly studied for Toeplitz matrices arising from scalar PDEs. Higher order discretization of PDEs and systems of PDEs naturally lead to block Toeplitz matrices where the individual blocks represent the dofs associated with one finite element or the coupling of the unknowns of different types. Recently, we started transferring the results for the scalar case to the case of systems that results in block-Toeplitz matrices or block-circulant matrices [Bolten, Donatelli, Ferrari and Furci, SIMAX 2022]. Besides studying higher-order discretizations of scalar PDEs, systems of PDEs also fit in this framework. Additionally, we considered systems with saddle point structure, applying recent results for multigrid for such systems [Notay, Numer. Math. 2016] to the structured matrix case [Bolten, Donatelli, Ferrari and Furci, LAA 2023; Bolten, Donatelli, Ferrari and Furci, APNUM 2023]. In the talk the analysis technique, the derived sufficient conditions for optimal convergence and numerical results will be presented.

Title: Symbol-Based Analysis of (Two Related) Multigrid Methods for Electromagnetic Scattering Problems

René Spoerer

Bergische Universität Wuppertal

Abstract:

The large null space of the curl operator presents a difficulty for standard multigrid approaches to certain classes of electromagnetic problems. As a remedy, the multigrid method developed by R. Hiptmair uses a hybrid two-step smoother to improve the convergence of the curl-free components. This talk presents a symbol-based spectral analysis of the system and iteration matrices involved, exploiting the circulant or Toeplitz structure that arises when the problem is discretized on a structured grid. We also compare the matrices and spectra with those arising in the closely related finite integration technique (FIT).

Title: Practical Advances in High-Order Stokes Solvers: Robust Multigrid Preconditioners with AMG Integration

Alexey Voronin

University of Illinois Urbana Champaign

Abstract:

This work introduces and assesses the efficiency of a novel monolithic phMG multigrid method, specifically designed for high-order discretizations of stationary Stokes systems using Taylor-Hood and Scott-Vogelius elements. The phMG approach integrates approximation order (p) and spatial (h) coarsening to address the computational and memory efficiency challenges that are often encountered in conventional high-order numerical simulations. Our comparative analysis reveals that phMG offers significant improvements over traditional spatial-coarsening-only multigrid (hMG) techniques for problems discretized with Taylor-Hood elements across a variety of problem sizes and discretization orders. In particular, the phMG method exhibits superior performance in reducing setup and solve times, particularly when dealing with higher discretization orders and unstructured problem domains. For Scott-Vogelius discretizations, while monolithic phMG delivers low iteration counts and competitive solve phase timings, it exhibits a discernibly slower setup phase when compared to multilevel full-block-factorization (FBF) preconditioners. This difference in efficiency stems from the incorporation of nested-mesh phMG into the FBF framework and lower setup costs for patch relaxation based on a singular unknown type, unlike monolithic phMG that requires the assembly of larger mixed-field relaxation patches, making the setup phase more costly in comparison.

CP 5. Domain decomposition preconditioners

Title: Brain edema simulation by using domain decomposition methods

Talal Alshehri

Morgan State University

Abstract:

In this paper, we consider using domain decomposition methods, particularly overlapping Schwarz preconditioners, to simulate brain edema using Biot's poroelasticity equations. Domain decomposition preconditioners which can solve in parallel are designed for reformulated Biot equations. Based on Schur complement theory, we developed Schur complement-based preconditioners and derived their approximate forms through Fourier analysis. We employ numerical experiments to validate the scalability of the proposed two-level overlapping Schwarz preconditioners. The results show that the number of iteration steps is determined by the overlapping ratio, which lays the foundation for brain simulation. We are conducting 2D simulations of brain edema to investigate the effects of physical parameters.

Title: Robust Domain Decomposition Methods for High-contrast Multiscale Problems on Irregular Domains

Juan Calvo

University of Costa Rica

Abstract:

We present a domain decomposition preconditioner for second-order elliptic partial differential equations that handles coefficients with high-contrast and multiscale properties, and is suitable for irregular subdomains. We will present partition of unity functions and appropriate eigenvalue problems that enrich usual coarse spaces. We demonstrate that the condition number of the preconditioned systems is bounded with a bound that is independent of the contrast, and include selected numerical experiments that confirm the robustness of our preconditioner.

References

- [1] Calvo, J. G., and Galvis, J. (2023). Robust domain decomposition methods for high-contrast multiscale problems on irregular domains with virtual element discretizations. *Journal of Computational Physics*, Volume 505, 2024.
- [2] Calvo, J. G. Virtual coarse spaces for irregular subdomain decompositions. In *Domain Decomposition Methods in Science and Engineering XXV*, Springer, 2020, 75–82.
- [3] Galvis, J., Chung, E. T., Efendiev, Y. and Leung, W. T. On overlapping domain decomposition methods for high-contrast multiscale problems. In *Domain Decomposition Methods in Science and Engineering XXIV*, Springer, 2018, 45–57.

Title: Preconditioned IDR Solution Methods in Scientific and Industrial Applications

Alex Fedoseyev

Ultra Quantum Inc.

Abstract:

A review of preconditioned solvers for large-scale applications in science and industry is presented. The analysis, parallelization, and optimization approach for large unstructured sparse matrices using IDR methods are considered for modern multicore microprocessors. CNSPACK is an advanced solver successfully used for the coupled solution of stiff problems arising in multiphysics applications, such as Computational Fluid Dynamics (CFD) for high Reynolds number turbulent flows, turbulent boundary layers, hypersonic and rarefied flows, carrier transport in semiconductors, and kinetic and quantum mechanics problems [2,3,4,5,6,7]. CNSPACK employs an iterative IDR algorithm with ILU preconditioning (where the user chooses the ILU order). Originally, CNSPACK was implemented and optimized for early sequential processors, considering their arithmetic and memory-size limitations. In the early 1990s, the first optimization exercise was performed to accelerate the algorithm by 6 times for emerging superscalar microprocessors, such as the Intel i860 [1]. However, there has been a significant shift in processor architectures and computer system organization since that time.

Nowadays, desktop computers and cluster nodes utilize high-performance pipelined superscalar multicore processors with out-of-order execution of instructions, deep cache hierarchies, and high-throughput memory capabilities. As a result, performance criteria and methods have been revisited, along with consideration of parallelization involving the solver and preconditioner using the OpenMP environment [8]. Results of the successful implementation for efficient parallelization are presented for computer systems based on Intel Core i7 or Xeon multicore multiprocessor architectures.

References

- [1] Fedoseyev A. , Bessonov O.(2001) Computational Fluid Dynamics Journal 10, 299-303.
- [2] Fedoseyev A., M. Turowski, L. Alles, and R. A. Weller (2008) Math. and Computers in Simulation 79, 1086-1096.
- [3] Fedoseyev A.I., Alexeev, B.V., Simulation of viscous flows with boundary layers within multiscale model using generalized hydrodynamics equations , Procedia Computer Science, 1 (2010) 665-672.
- [4] Fedoseyev A., Alexeev B.V., Generalized hydrodynamic equations for viscous flows-simulation versus experimental data, in AMiTaNS-12, American Institute of Physics AIP CP 1487, 2012, pp.241-247.
- [5] Fedoseyev A., Finite element method stabilization for supersonic flows with flux correction transport method, AIP CP 2302, 120003 (2020), Ed. M.Todorov,.
- [6] Fedoseyev A., Griaznov V., Simulation of Rarefied Hypersonic Gas Flow and Comparison with Experimental Data, in AMiTaNS-2021, Conf. Proc., AIP CP 2522, 100003, 2021, Ed. M.Todorov. AmiTaNS'23 Journal of Physics: Conference Series 2675 (2023) 012011, IOP Publishing
- [7] Fedoseyev A., Griaznov V., Ouazzani J., Simulation of rarefied hypersonic gas flow and comparison with experimental data II, Proc. AMITANS-2022 Conf., AIP CP 2953, 2023, Ed. M.Todorov.
- [8] Bessonov O. A. , Fedoseyev A., Parallelization of the Preconditioned IDR Solver for Modern Multicore Computer Systems, Proc. AMITANS-2012 Conf., American Institute of Physics, AIP CP 1487, 2012, 314-321, Ed. M.Todorov.

CP 6. Advanced Preconditioning Techniques

Title: Preconditioning Techniques for Multiterm Generalized Sylvester Equations

Yannis Voet

École Polytechnique Fédérale de Lausanne

Abstract:

Sylvester matrix equations are ubiquitous in scientific computing. However, few solution techniques exist for their generalized multiterm version, as they now arise in an increasingly large number of applications. In this talk, I present two algebraic parameter free preconditioning techniques for iteratively solving multiterm Sylvester equations. They consist in either constructing a low Kronecker rank approximation of the operator itself or its inverse. While

applying the preconditioning operator for the former requires solving a standard Sylvester equation at each iteration, the latter only requires matrix-matrix multiplications, which are highly optimized on modern computer architectures. Moreover, low Kronecker rank approximate inverses can be easily combined with sparse approximate inverse techniques, thereby further speeding up their application without adversely impacting their effectiveness. Finally, the methods are tested for various applications.

Title: Preconditioning for Topological Constraint Problem

Mingdong He

University of Oxford

Abstract:

The Parker problem remains an open question since it was first proposed in 1972 [1]. It states that given a magnetic configuration, the static equilibrium magnetic field will be tangentially discontinuous. This relaxation process can be described by ideal magnetohydrodynamics equations. For such a time-dependent problem, one of the challenges lies in numerical methods which can preserve the topology of the magnetic field [2], called helicity, which acts as a topological barrier for the energy decay [3]. At the same time investigating the static solution requires a fast and robust solver concerning the physical parameters in the PDEs, which control the energy decay. In this talk, we will first outline the numerical approaches to investigate the Parker problem, which includes finite element structure-preserving discretisation, initial condition truncation and magnetic potential computation for helicity. Then we will present the unsymmetric structure of the Newton linearisation but with a nice Schur complement. Finally, we will discuss the physical parameters and the preconditioning techniques to achieve parameter robustness and the challenges. Numerical results will be shown.

References

- [1] Eugene N. Parker. Topological dissipation and the small-scale fields in turbulent gases. *Astrophysical Journal*, vol. 174, p. 499, 174:499, 1972.
- [2] David I. Pontin and Gunnar Hornig. The Parker problem: Existence of smooth force-free fields and coronal heating. *Living Reviews in Solar Physics*, 17(1):5, December 2020.
- [3] Boris Khesin. *Topological Fluid Dynamics*. 52(1), 2005.

Title: Data-Driven Solver and Preconditioner Selection for Sparse Linear Matrices

Hayden Liu Weng

Technical University of Munich

Abstract:

Solving large, sparse linear systems is at the core of diverse computational domains, where the efficient solution of such systems can heavily impact the total execution time of computations. While applying a preconditioner to an iterative solver has become standard, making optimal or

sometimes even numerically stable choices can be quite challenging, mainly because the best combination depends strongly on the specific problem. We discuss how to predict effective preconditioner and iterative solver combinations for any given sparse linear system using a data-driven approach based on a combination of embedding and linear modeling techniques. We focus on determining useful system features and investigate different metrics to quantify the relative performance of the preconditioned solvers across matrices from the SuiteSparse collection.

CP 7. Advances in Multigrid Preconditioners

Title: Nesting Approximate Inverses for Improved Preconditioning and Algebraic Multigrid Smoothing

Andrea Franceschini

University of Padova

Abstract:

Approximate inverses are a highly valuable tool for preconditioning and algebraic multigrid (AMG) smoothing due to their high degree of parallelism, making them ideal for exploiting high-performance computing environments. However, one of the most limiting drawbacks is the rapid increase in setup costs as density increases, thereby restricting the ability to improve accuracy simply by adding more entries. In this study, we examine the use of approximate inverses in factored form and emphasize the significant improvement in effectiveness that can be achieved by nesting more factors while keeping computational costs reasonable. Additionally, we suggest strategies and offer theoretical insights to lessen the computational overhead associated with the triple matrix product required during initial nesting stages. Numerical experiments conducted across a range of real-world applications demonstrate the efficacy and effectiveness of our proposed approach.

Title: LFA-tuned matrix-free multigrid for the elastic Helmholtz equation

Rachel Yovel

Ben-Gurion University of the Negev

Abstract:

The Helmholtz equation arises in modeling wave propagation in the frequency domain. The acoustic Helmholtz equation models acoustics and electromagnetics, while the elastic Helmholtz equation models wave propagation in solids, such as the earth's sub-surface. Both are difficult to solve numerically, as the discrete linear system is very large, indefinite, and ill-conditioned. The elastic version amplifies these difficulties both because of its larger size (as a system of PDEs) and its more complicated physics. We present an efficient matrix-free geometric multigrid method for the elastic Helmholtz equation, and a suitable discretization. Many discretization methods had been considered in the literature for the Helmholtz equations,

as well as many solvers and preconditioners, some of which are adapted for the elastic version of the equation. However, there is very little work considering the reciprocity of discretization and a solver. We take steps towards bridging this gap, as our discretization is chosen to fit the solver. Our multigrid method is based on the shifted Laplacian approach, together with approaches used for linear elasticity. Our discretization for the elastic Helmholtz equation is inspired by an existing fourth-order stencil for acoustic Helmholtz. Using two-grid local Fourier analysis, we validate the compatibility of our discretization with our solver and tune a choice of weights for the stencil that optimize the convergence rate of the multigrid cycle. The resulting discretization reduces numerical dispersion, and hence improves the coarse grid correction. We show, numerically and theoretically, that our discretization allows the use of less grid points per shear wavelength without deteriorating the performance. It results in a scalable multigrid preconditioner that can tackle large real-world 3D scenarios.

Title: Bi-parametric Operator Preconditioning

Carlos Jerez-Hanckes

University of Bath, Universidad Adolfo Ibáñez

Abstract:

We extend the operator preconditioning framework Hiptmair (2006) [10] to Petrov-Galerkin methods while accounting for parameter-dependent perturbations of both variational forms and their preconditioners, as occurs when performing numerical approximations. By considering different perturbation parameters for the original form and its preconditioner, our bi-parametric abstract setting leads to robust and controlled schemes. For Hilbert spaces, we derive exhaustive linear and super-linear convergence estimates for iterative solvers, such as h -independent convergence bounds, when preconditioning with low-accuracy or, equivalently, with highly compressed approximations.

Minisymposium

MS 1. Parallel and Machine Learning Preconditioning Methods for Large Linear Systems

Title: Prospective on Latest Advances of Scalable Hybrid Monte Carlo Methods for Linear Algebra

Vassil Alexandrov

Hartree Centre

Abstract:

This paper provides some results of our current investigation of the applicability of hybrid Monte Carlo methods for solving systems of linear algebraic equations to a variety of problems in science and engineering. In particular Markov Chain Monte Carlo Matrix Inversion (MCMCMI) is used as a preconditioner in combination with GMRES and (Bi)CG(stab) methods to solve a variety of problems arising in quantum chromodynamics, plasma physics and engineering. Representative matrices for the latter two are extracted from BOUT++ and Nektar++ implementations of specific simulation scenarios.

The results on the performance and scalability of the implementations of the method in C++/CUDA and Python/CuPy for a variety of CPU and GPU architectures (e.g., P100, V100, A100), as well as, our preliminary observations on the effects of using QRNG (quantum random number generator) will be presented.

Title: Matrix-Free Parallel Scalable Multilevel Deflation Preconditioning for the Helmholtz Equation

Jinqiang Chen

TU Delft

Abstract:

We present a matrix-free parallel scalable multilevel deflation preconditioned method for heterogeneous time-harmonic wave problems, including Helmholtz and elastic wave equations. Building upon recent advances in deflation preconditioning [1, 2] for highly indefinite time-harmonic waves, we adapt these techniques for parallel implementation in the context of solving large-scale heterogeneous problems with minimal pollution error. The proposed method integrates the Complex Shifted Laplacian preconditioner (CSLP) with deflation approaches, employing higher-order deflation vectors and re-discretization schemes derived from the Galerkin coarsening approach for a matrix-free parallel implementation. We suggest a robust and efficient configuration of the matrix-free multilevel deflation method, which yields a near-wavenumber-independent convergence and improved time efficiency. Numerical experiments

demonstrate the effectiveness of our approach for increasingly complex model problems. The matrix-free implementation of the preconditioned Krylov subspace methods reduces memory consumption, and the parallel framework exhibits satisfactory parallel performance. This work represents a significant step towards developing efficient, scalable, and parallel multilevel deflation preconditioning methods for large-scale real-world applications in wave propagation.

References

- [1] V. Dwarka, C. Vuik (2020). Scalable convergence using two-level deflation preconditioning for the Helmholtz equation. *SIAM Journal on Scientific Computing*, 42(2), A901–A928.
- [2] V. Dwarka, C. Vuik (2022). Scalable multi-level deflation preconditioning for highly indefinite time-harmonic waves. *Journal of Computational Physics*, 469, 111327.

Title: Scalable distributed preconditioners in Ginkgo

Pratik Nayak

Karlsruhe Institute of Technology

Abstract:

Efficient preconditioners are critical in scientific applications to accelerate the solution of linear systems. The latest exascale machines, being heterogeneous and GPU-centric, require implementations that ensure efficient utilization of fine-grained parallelism of the GPUs, while distributing them over multiple of these heterogeneous nodes. In this talk, we will present our approach to performance-portable distributed preconditioners in Ginkgo. We will talk about our approach to distributed multigrid, which separates the coarsening methods from the multigrid level hierarchy enabling users to compose different smoothers, coarsening methods, coarse solvers and precisions in a very efficient manner. We will also briefly talk about domain decomposition preconditioners in Ginkgo such as Schwarz and balanced domain decomposition by constraints (BDDC). Finally, we will show performance results for these preconditioners and some examples from applications which use them.

Title: DeepONet-based Preconditioning for Krylov Methods

Alena Kopanicakova

Università della Svizzera italiana

Abstract:

We introduce a new class of hybrid preconditioners for solving parametric linear systems of equations. The proposed preconditioners are constructed by hybridizing the deep operator network, namely DeepONet, with standard iterative methods. Exploiting the spectral bias, DeepONet-based components are harnessed to address low-frequency error components, while conventional iterative methods are employed to mitigate high-frequency error components. Our preconditioning framework utilizes the basis functions extracted from pre-trained DeepONet to construct a map to a smaller subspace, in which the low-frequency

component of the error can be effectively eliminated. Our numerical results demonstrate that the proposed approach enhances the convergence of Krylov methods by a large margin compared to standard non-hybrid preconditioning strategies. Moreover, the proposed hybrid preconditioners exhibit robustness across a wide range of model parameters and problem resolutions.

MS 2. Preconditioners for High-Frequency Helmholtz Problems

Title: Towards scalable preconditioners for indefinite systems arising in electromagnetic simulations

Vandana Dwarka

TU Delft

Abstract:

Many applications ranging from imaging to the design of nuclear fusion devices rely on solving indefinite linear systems arising from certain partial differential equations. While state-of-the-art solvers exist for symmetric and positive-definite systems, nonsymmetric indefinite problems, such as the Helmholtz problem, remain notoriously difficult to solve numerically in the absence of scalable solvers. In this talk, theory and algorithms for this highly indefinite problem will be introduced, as well as extensions to serve as a baseline for other PDEs leading to indefinite systems. In particular, we will discuss recent developments of the scalable deflation preconditioner and multigrid as a stand-alone solver for this long-standing open problem.

Title: Using Spectral Coarse Spaces of the H-GenEO Type for Efficient Solutions of the Helmholtz Equation

Victorita Dolean

TU Eindhoven

Abstract:

The Helmholtz equation is a widely used model in wave propagation and scattering problems. However, its numerical solution can be computationally expensive in high-frequency regime due to the oscillatory solution and the potential contrasts in coefficients. Parallel domain decomposition methods have been identified as promising solvers for such problems, but they often require a suitable coarse space to achieve robust behaviour. In this talk, we present the H-GenEO coarse space, which constructs an effective coarse space using localized eigenvectors of the Helmholtz operator. While the GenEO coarse space is designed for symmetric positive definite problems, the theory cannot be extended directly to the H-GenEO coarse space due to the indefinite nature of the underlying problem. During this talk it will be shown what the H-GenEO coarse space is capable of providing the required robust behaviour when used with a

suitable domain decomposition method. Numerical experiments for increasing wave numbers demonstrate the efficiency of the method in solving complex Helmholtz problems, with potential applications in various scientific and engineering domains.

Title: Acceleration of non-local exchange in generalized optimized Schwarz methods

Xavier Claeys

Sorbonne Université

Abstract:

The generalized optimized Schwarz method proposed in [Claeys & Parolin, 2022] is a variant of the Després algorithm for solving harmonic wave problems where transmission conditions are enforced by means of a non-local exchange operator. Compared to the original Després algorithm where transmission conditions are expressed in terms of a simple swap of unknowns, which is easy to compute, this novel exchange operator induces a non-negligible additional computational cost. I shall present an easily implementable acceleration technique that significantly reduces this cost without any deterioration on the precision and convergence speed of the overall domain decomposition algorithm. It combines both preconditioning and recycling techniques. I will present numerical experiments and theoretical estimates. This is joint work with Roxane Atchekzai (SU/CEA) and Matthieu Lecouvez (CEA).

Title: Some convergence results for RAS-Imp and RAS-PML for the Helmholtz equation

Shihua Gong

The Chinese University of Hong Kong, Shenzhen, and SICIAM, SRIBD

Abstract:

We consider two variants of restricted overlapping Schwarz methods for the Helmholtz equation. The first method, known as RAS-Imp, incorporates impedance boundary condition to formulate the local problems. The second method, RAS-PML, employs local perfectly matched layers (PML). These methods combine the local solutions additively with a partition of unity. We have shown that RAS-Imp has power contractivity for strip domain decompositions. More recently, we showed that RAS-PML has super-algebraic convergence with respect to wavenumber after a specified number of iterations. This is the first theoretical result for the nontrapping Helmholtz problems with variable wave speed. In this talk we review these results and then investigate their sharpness using numerical experiments. We also investigate situations not covered by the theory. In particular, the theory needs the overlap of the domains or the PML widths to be independent of k . We present numerical experiments where these distances decrease with k .

MS 3. Recent advances in multigrid preconditioning

Title: An interior-point multigrid-based approach for scalable computational contact mechanics

Tucker Hartland

Lawrence Livermore National Laboratory

Abstract:

A critical aspect of modeling complex engineering systems is the interaction of physical bodies in contact. A number of frictionless contact problems have the property that the modeled state is the minimizer of an energy objective functional. However, generally such problems are: nonlinear, nonconvex and contain an optimization variable whose dimension is unbounded with respect to mesh refinement. We focus on the scalable solution of such large-scale contact mechanics problems on high-performance computing systems. We employ a Newton-based interior-point filter line-search method, that has emerged as one of the most robust methods for nonlinear nonconvex constrained optimization, to computationally estimate minimizers of such large-scale constrained optimization problems. The outer Newton-based interior-point loop converges rapidly; however, each step requires the solution of a large saddle-point linear system. A major challenge with the inner interior-point Newton-based linear system is that, in addition to the general challenges of solving large-scale linear systems, it can become arbitrarily ill-conditioned as the optimizer estimate approaches the optimal point. There are blocks that are amenable to multigrid which is a problem feature that we exploit. In this talk, we detail an interior-point multigrid-based approach for solving such problems and present scaling results obtained from an implementation of said framework on a few contact mechanics example problems. The results show that the solution of various contact mechanics problems can be achieved in a manner that scales well in the large-scale regime.

Title: Robust physics-based preconditioners for multi-physics problems

Xiaozhe Hu

Tufts University

Abstract:

We are interested in reliable simulations of some biophysical processes in the brain, such as blood flow and metabolic waste clearance. Modeling those processes results in interface-driven multi-physics problems that can be coupled across dimensions. However, the complexity of the interface coupling often deteriorates the performance of standard methods in finding the numerical solution. Therefore, based on the physics-based operator preconditioning framework, we design parameter-robust preconditioners that specifically target such multi-physics problems. Different techniques, such as rational approximations for fractional operators and specially tailored algebraic multigrid method that preserves the coupling information, are developed to handle the coupling that enforces interface constraints. We theoretically prove the parameter-robustness of the proposed preconditioners. We also present several numerical examples of realistic geometries, such as the viscous-porous flow coupling of the cerebrospinal

fluid or the mixed-dimensional model of flow in vascularized brain tissue, to demonstrate their effectiveness and scalability in practical applications.

Title: A multigrid reduction framework for multi-physics applications

Victor Magri

Lawrence Livermore National Laboratory

Abstract:

Solving multiphysics problems, such as subsurface fluid flow, involves tightly coupled systems that present significant challenges due to large, non-symmetric, and ill-conditioned linear systems. While fully implicit methods are effective, they require robust preconditioners for fast convergence. At the same time, the advent of HPC hardware with GPUs offered significant performance improvements compared to CPU-based supercomputers. This work presents recent advances to the multigrid reduction (MGR) framework, a general preconditioner for the solution of multiphysics problems implemented in hypr. We show that MGR is flexible to accommodate a wide range of scenarios, and we demonstrate its applicability and scalability for real-world subsurface flow simulations on modern HPC systems, including the fastest supercomputer in the world as of May 2024.

Title: Mixed precision algorithm development in hypr

Ulrike Yang

Lawrence Livermore National Laboratory

Abstract:

The hypr software library provides parallel solvers and preconditioners for a variety of high-performance computing architectures. Recently, the use of mixed precision in algorithms has become of high interest since it provides reduced memory use and faster performance for lower precision operations. Development of hypr started more than twenty-five years ago, and generally focused on using double precision. While hypr can also be configured at single precision, it was not originally designed to allow the use of several different precisions in combination. Recently, the hypr team developed the capability of using mixed precision in hypr. This talk will describe the new capability including the challenges that needed to be overcome and present some results of mixed precision algebraic multigrid methods.

MS 4. Analog and mixed precision preconditioning

Title: Solvers and Preconditioners for Analog Architectures

Erik Boman

Sandia National Laboratories

Abstract:

As Moore's law is stalling, new architectures are needed to ensure ever higher compute power. A promising technology is analog "in situ" systems, such as memristive crossbars. These systems provide significant speedup for dense matrix-vector multiplication, at a fraction of the power. We discuss how a hybrid analog-digital system can be used to solve linear systems using preconditioned Krylov methods. This is closely related to mixed-precision computing, where the analog part provides low precision and the digital part high precision.

Title: Solving Sparse Linear Systems via Flexible GMRES with In-Memory Analog Preconditioning

Chai Wah Wu

IBM Research

Abstract:

Analog arrays of non-volatile crossbars leverage physics to compute approximate matrix-vector multiplications in a rapid, in-memory fashion. In this paper we consider exploiting this technology to precondition the Generalized Minimum Residual iterative solver (GMRES). Since the preconditioner must be applied through matrix-vector multiplication, approximate inverse preconditioners are a natural fit. At the same time, the errors introduced by the analog hardware render an iteration matrix that changes from one iteration to another. To remedy this, we propose to combine analog approximate inverse preconditioning with a flexible GMRES algorithm that naturally incorporates variations of the preconditioner into its model. The benefit of our approach is that the analog circuit is much simpler than correcting the errors at the hardware level. Our experiments with a simulator for analog hardware show that such an analog-flexible scheme can lead to fast convergence.

Title: Half precision wave simulation

Longfei Gao

Argonne National Laboratory

Abstract:

On modern hardware, the speed of memory access is often the limiting factor for execution time for many scientific and industrial applications, particularly for those involving PDE discretizations that exploit sparsity. This motivates us to explore the possibility of operating at half precision to reduce memory footprint and hence utilize the memory bandwidth more effectively. We study the viability of half precision simulations for time dependent wave equations. Potential pitfalls of naively switching to half precision will be illustrated, including nonphysical oscillations and energy loss. An effective remedy in the form of compensated sum will be presented, which is able to restore the simulation quality to a satisfactory level, as illustrated by numerical examples on modern GPUs.

MS 5. Recent Advances in Saddle-Point and Double Saddle-Point Systems

Title: Spectral Properties of Double Saddle-Point Systems

Chen Greif

The University of British Columbia

Abstract:

Double saddle-point systems are drawing increasing attention in the past few years, due to the importance of multiphysics and other relevant applications and the challenge in developing efficient iterative numerical solvers. In this talk we describe some of the numerical properties of the matrices arising from these problems. We discuss invertibility conditions, derive eigenvalue bounds, and show that if Schur complements are effectively approximated, the eigenvalue structure gives rise to rapid convergence of Krylov subspace solvers. A few numerical experiments illustrate our findings.

Title: Block Triangular Preconditioners for Double Saddle-Point Problems Arising in Mixed Hybrid Coupled Poromechanics

Massimiliano Ferronato

University of Padova

Abstract:

In this communication, we describe and analyze the spectral properties of a class of inexact block triangular preconditioners for double saddle-point symmetric linear systems arising from the mixed finite element and mixed hybrid finite element discretization of Biot's poroelasticity equations. We develop a spectral analysis of the preconditioned matrix, showing that the complex eigenvalues lie in a circle of center $(1,0)$ and radius at most 1, while the real eigenvalues are described in terms of the roots of a third order polynomial with real coefficients. Numerical examples are reported to verify the quality of the theoretical bounds and illustrate the efficiency of the inexact versions of the proposed preconditioners, especially in comparison with similar block diagonal strategies along with the MINRES iteration.

Title: An Augmented Lagrangian Preconditioner for the Control of the Navier--Stokes Equations

Santolo Leveque

Scuola Normale Superiore di Pisa

Abstract:

Optimal control problems with PDEs as constraints arise very often in scientific and industrial applications. Due to the difficulties arising in their numerical solution, researchers have put a great effort into devising robust solvers for this class of problems. An example of a highly challenging problem attracting significant attention is the distributed control of incompressible viscous fluid flow problems. In this case, the physics is described by the incompressible Navier--Stokes equations. Since the PDEs given in the constraints are non-linear, in order to obtain a solution of Navier--Stokes control problems one has to iteratively solve linearizations of the problems until a prescribed tolerance on the non-linear residual is achieved. In this talk, we present efficient and robust preconditioned iterative methods for the solution of the stationary incompressible Navier--Stokes control problem, when employing a Gauss--Newton linearization of the first-order optimality conditions. The iterative solver is based on an augmented Lagrangian preconditioner. By employing saddle-point theory, we derive suitable approximations of the $(1,1)$ -block and the Schur complement. Numerical experiments show the effectiveness and robustness of our approach, for a range of problem parameters.

This is joint work with Michele Benzi (Scuola Normale Superiore) and Patrick Farrell (University of Oxford).

MS 6. Nonlinear Preconditioning Techniques and Applications I

Title: Some preconditioned inexact Newton methods with learning capabilities

Xiao-Chuan Cai

University of Macau

Abstract:

We discuss a stagnation shortening inexact Newton algorithm with a learning phase during the Newton iterations in which the residual subspace is centralized and decomposed into a slow subspace and a regular subspace using an unsupervised learning method based on the principal component analysis. We show numerically that with such an embedded learning phase the inexact Newton method converges almost quadratically. As an application, we consider the modeling of the human artery with stenosis using the hyperelasticity equation with multiple material parameters. Due to the significant difference in the material coefficients between the plaques and the healthy parts of the blood vessels, the problem is nonlinearly very difficult. Numerical experiments demonstrate that proposed method offers significantly reduced number of nonlinear iterations and robustness. This is a joint work with L. Luo and Y. Gong.

Title: Domain decomposition preconditioners and multi-scale approaches to solve stationary and time-dependent nonlinear equations

Victorita Dolean

TU Eindhoven

Abstract:

In this contribution, we extend a previous work done by the authors, in which they introduced a coarse space for the Poisson equation posed on the perforated domains containing multiscale features, as they arise in simplified flow models in an urban environment. Here, the focus is on left nonlinear preconditioning techniques based on overlapping subdomains, implementing techniques that use the coarse space proposed in the linear case to provide scalability. The coarse space was used in combination with the RAS preconditioner, an overlapping domain decomposition technique for the solution of linear problems. Here we compare numerically different preconditioning strategies for a given model problem. While the coarse space was originally based on the linear Poisson equation, we find that it is a fitting coarse space for nonlinear problems as well.

Title: Nonlinear Preconditioning for Implicit Solution of Discretized PDEs

David Keyes

King Abdullah University of Science and Technology

Abstract:

Nonlinear preconditioning refers to transforming a nonlinear algebraic system to a form for which Newton-type algorithms have improved success through quicker advance to the domain of quadratic convergence. We place these methods in the context of a proliferation of variations distinguished by being left- or right-sided, multiplicative or additive, non-overlapping or overlapping, and partitioned by field, subdomain, or other criteria. We present the Nonlinear Elimination Preconditioned Inexact Newton, which is based on a heuristic bad/good heuristic splitting of equations and corresponding degrees of freedom. We augment basic forms of nonlinear preconditioning with three features of practical interest: a cascadic identification of the bad discrete equation set, an adaptive switchover to ordinary Newton as the domain of convergence is approached, and error bounds on output functionals of the solution. Various nonlinearly stiff algebraic and model PDE problems are considered for insight and we illustrate performance advantage and scaling potential on challenging two-phase flows in porous media.

MS 7. Preconditioned Linear Algebraic Techniques for Solving Inverse Problems

Title: Effective Approximate Preconditioners for Linear Inverse Problems

Lucas Onisk

Emory University

Abstract:

Many problems in science and engineering give rise to linear systems of equations that are commonly referred to as large-scale linear discrete ill-posed problems. These problems arise for instance, from the discretization of Fredholm integral equations of the first kind. The matrices that define these problems are typically severely ill-conditioned and may be rank deficient. Because of this, the solution of linear discrete ill-posed problems may not exist or be extremely sensitive to perturbations caused by error in the available data. These difficulties can be reduced by applying regularization to iterative refinement type methods which may be viewed as a preconditioned Landweber method. Using a filter factor analysis, we demonstrate that low precision matrix approximants can be useful in the construction of these preconditioners.

Title: A New Deflation Space for Preconditioned GMRES

Daniel Szyld

Temple University

Abstract:

New convergence bounds are presented for weighted, preconditioned, and deflated GMRES for the solution of large, sparse, nonsymmetric linear systems, where it is assumed that the symmetric part of the coefficient matrix is positive definite. The new bounds are sufficiently explicit to indicate how to choose the preconditioner and the deflation space to accelerate the convergence. One such choice of deflating space is presented, and numerical experiments illustrate the effectiveness of such space. Joint work with Nicole Spillane (Ecole Polytechnique).

Title: Preconditioning Linear Inverse Problems Using Randomization and Subspace Projection

Eric de Sturler

Virginia Tech

Abstract:

Title: Randomized Approaches for Optimal Experiment Design

Srinivas Eswar

Argonne National Laboratory

Abstract:

This talk is regarding linear systems that arise in Bayesian linear inverse problems. The first part is on the efficient construction of scalable preconditioners of the Gauss-Newton Hessian. The second part is on using the prior-preconditioned forward operator to inform sensor placement decisions in optimal experiment design. Both approaches uses recent advances in randomized numerical linear algebra and come with strong theoretical guarantees. Numerical experiments on model inverse problems demonstrate the

effectiveness of these methods. This is joint work with Amit Subrahmanya, Vishwas Rao, Arvind K. Saibaba.

MS 8. Algebraic and Geometric Domain Decomposition Preconditioners for Complex Problems

Title: Substructuring the Hiptmair-Xu Preconditioner

Xavier Claeys

Sorbonne Université

Abstract:

Considering positive Maxwell problems in 3D discretized by low order Nédélec edge elements, we propose a substructured variant of the Hiptmair-Xu preconditioner based on a new formula that expresses the inverse of Schur systems in terms of the inverse matrix of the global volume problem. We obtain condition number estimates stemming from those available for the original Hiptmair-Xu preconditioner. Besides theory, we shall present numerical results confirming stabilisation of the condition number with respect to the meshwidth.

Title: Overlapping Schwarz Preconditioner with Geneo Coarse Space for Nonlocal Equations

Pierre Marchand

INRIA Paris Saclay

Abstract:

Domain Decomposition Methods, such as Additive Schwarz, can be used to precondition linear systems, and they usually rely on an additional coarse space to scale with the number of subdomains. The Generalized Eigenproblems in the Overlaps (GenEO) has emerged as one of the most promising coarse space for sparse symmetric positive definite problems, see Spillane et al. (2014). GenEO takes eigenvectors of well-chosen local eigenproblems as a basis for the coarse space. As one of its interesting features, GenEO is only based on the knowledge of the stiffness matrix elements and discretization agnostic, left apart a few reasonable assumptions. Recently, the GenEO approach has been extended to Boundary Integral Equations (BIEs) for the hypersingular operator in Marchand et al. (2020). In this context, the discretized operator is non-local so that the resulting linear system is dense. Thus, the local eigenproblems used to build the GenEO coarse space are adapted to the non-local nature of the problem and its $(\|\cdot\|_{\widetilde{H}^{1/2}})$ energy norm. In this talk, we will present theoretical and numerical results aiming at adapting GenEO to the integral fractional Laplacian of order $(s \in (0,1))$. It shares many similarities with BIEs, e.g. its non-local nature and the energy norm $(\|\cdot\|_{\widetilde{H}^s})$ that will be used to introduce a new distributed solver using the libraries PyNucleus, Htool-DDM and HPDDM.

Title: An Algebraic Domain Decomposition Preconditioner

Nicole Spillane

CNRS, Ecole Polytechnique

Abstract:

Domain decomposition are an efficient class of preconditioners for solving large scale problems on parallel computers. A crucial step is choosing the deflation, or coarse, space. In this talk a coarse space is introduced for symmetric positive definite linear systems. It is called AWG (for Algebraic-Woodbury-GenEO) and constructed algebraically: only the knowledge of the matrix A for which the linear system is being solved is required. Thanks to the GenEO spectral coarse space technique, the condition number of the preconditioned operator is bounded theoretically from above. This upper bound can be made smaller by enriching the coarse space with more spectral modes. The novelty is that, unlike in previous work on the GenEO coarse spaces, no knowledge of a partially non-assembled form of A is required. Indeed, the spectral coarse space technique is not applied directly to A but to a low-rank modification of A of which a suitable nonassembled form is known by construction. The extra cost is a second (and to this day rather expensive) coarse solve in the preconditioner.

Title: Development of preconditioning techniques for integrated energy systems

Buu-Van Nguyen

Delft University of Technology

Abstract:

With the ongoing energy transition, more sustainable energy sources are required. This increases interaction amongst single-carrier energy systems such as gas, electricity and heat. These interacting systems are called integrated energy systems. The energy transport capabilities are of interest for planning and operating purposes. Modelling the load flow of integrated energy systems results in a nonlinear system. The Newton-Raphson method is a common way to solve this system, which leads to a sparse Jacobian. Our ambition is to model on the scale of the European energy system including gas, electricity and heat. This results in a system of size $n > 10^9$. Thus Krylov solvers with a scaleable preconditioner are particularly attractive for this problem

MS 9. Nonlinear Preconditioning Techniques and Applications I

Title: Batch Normalization Preconditioning for Neural Network Training

Qiang Ye

University of Kentucky

Abstract:

Batch normalization (BN) is a popular and ubiquitous method in deep neural network training that has been shown to decrease training time and improve generalization performance. Despite its success, BN is not theoretically well understood. It is not suitable for use with very small mini-batch sizes or online learning. In this talk, we will review BN and present a preconditioning method called Batch Normalization Preconditioning (BNP) to accelerate neural network training. We will analyze the effects of mini-batch statistics of a hidden variable on the Hessian matrix of a loss function and propose a parameter transformation that is equivalent to normalizing the hidden variables to improve the conditioning of the Hessian. Compared with BN, one benefit of BNP is that it is not constrained on the mini-batch size and works in the online learning setting. We will present several experiments demonstrating competitiveness of BNP.

Title: Batch Normalization Preconditioning for Convolutional Neural Networks

Susanna Lange

University of Chicago

Abstract:

In this talk we explore a new method of preconditioning applied during neural network training called Batch Normalization Preconditioning (BNP). Instead of applying normalization explicitly through a batch normalization layer as is done in Batch normalization (BN), BNP applies normalization by conditioning the parameter gradients directly during training. This is designed to improve the Hessian matrix of the loss function and hence convergence during training. In this talk, we consider how BNP applies to Convolutional Neural Networks (CNNs) by deriving the preconditioning matrix for CNNs. Furthermore, we explore how this derivation provides a theoretical understanding of how BN should be applied to Convolutional Neural Networks.

Title: A Gromov--Wasserstein Geometric Objective for Graph Coarsening and Potentials for Preconditioning

Jie Chen

MIT-IBM Watson AI Lab, IBM Research

Abstract:

Graph coarsening is a technique for solving large-scale graph problems by working on a smaller version of the original graph, and possibly interpolating the results back to the original graph. Popularized by algebraic multigrid methods applied to solving linear systems of equations, graph coarsening finds a new chapter in machine learning, particularly graph-based learning models. However, it is challenging to naively apply existing coarsening methods, because it is unclear how the multigrid intuition matches the machine learning problem at hand. We develop an objective-driven approach by explicitly defining the coarsening objective, which admits a

geometric interpretation—maintaining the pairwise distance of graphs. We derive the objective function by bounding the change of the distance and show its relationship with weighted kernel k-means clustering, which subsequently defines the coarsening method. We demonstrate its effective use in graph regression and classification tasks.

Title: A structure-guided Gauss-Newton method for shallow ReLU neural network

Tong Ding

Purdue University

Abstract:

In this talk, we propose a structure-guided Gauss-Newton (SgGN) method for solving least squares problems using a shallow ReLU neural network. By categorizing the weights/bias of the hidden and output layers of the network as nonlinear and linear parameters, the method iterates back and forth between the nonlinear and linear parameters. The nonlinear parameters are updated by a damped Gauss-Newton method, and the linear ones are updated by a linear solver. Moreover, at the Gauss-Newton step, a special form of the Gauss-Newton matrix is derived for the shallow ReLU neural network and is used for efficient computations. It is shown that the corresponding mass and Gauss-Newton matrices in the respective linear and nonlinear steps are symmetric and positive definite under reasonable assumptions. The SgGN method was tested for several one and two dimensional least-squares problems which are difficult for commonly used training algorithms in machine learning such as BFGS and ADAM. The loss curves for all four test problems clearly show that the SgGN out-performs those methods by a very large margin. This conclusion is further enhanced by examining ability and efficiency of the methods on moving the breaking hyper-planes (points for one dimension and lines for two dimensions). The breaking hyper-planes are determined by the nonlinear parameters (weights and bias of the hidden layer).

MS 10. Preconditioning and Machine Learning II

Title: Equivariant Generative Models for Molecular Modeling

Bao Wang

University of Utah

Abstract:

Molecular modeling tasks exhibit different symmetries, e.g. roto-translation equivariance and periodicity. A grand challenge in machine learning-assisted molecular modeling -- e.g. molecule generation -- is to account for different inherent symmetries. In this talk, I will discuss a few issues on building and training stable and expressive equivariant generative models, including normalizing flows and diffusion models, for molecule generations. Furthermore, I will discuss the role of steerable features of different types for equivariant machine learning.

Title: Generating Polynomial Method for Non-symmetric Tensor Decomposition

Zequn Zheng

Louisiana State University

Abstract:

Tensors or multidimensional arrays are higher-order generalizations of matrices. They are natural structures for expressing data that have inherent higher-order structures. Tensor decompositions play an important role in learning those hidden structures. In this talk, we present a novel algorithm to find the tensor decompositions utilizing generating polynomials. Under some conditions on the tensor's rank, we prove that the exact tensor decomposition can be found by our algorithm. Numerical examples successfully demonstrate the robustness and efficiency of our algorithm.

Title: Fast solvers for neural network least-squares approximations

Jianlin Xia

Purdue University

Abstract:

Neural networks provide an effective way to approximate functions, especially for some challenging situations with discontinuities, large variations, and sharp transitions. In our recent development of a novel block Gauss-Newton method for least-squares approximations via ReLU shallow neural networks, some dense linear systems arise in the iterations for finding some linear and nonlinear parameters. The coefficient matrices are shown to be symmetric and positive definite. We can further show that they are highly ill conditioned, and the condition numbers get even worse for some challenging function approximations. The ill-conditioned dense linear systems are thus difficult to solve by traditional direct and iterative solvers. On the other hand, we prove that the matrices have some interesting features that we can explore so that the systems can be solved efficiently and accurately. This is joint work with Zhiqiang Cai, Tong Ding, Min Liu, and Xinyu Liu.

MS 11. Nonlinear Preconditioning Techniques and Applications II

Title: Exploring nonlinear preconditioning strategies for solving phase-field fracture problems

Hardik Kothari

Università della Svizzera italiana

Abstract:

The phase-field approach has gained significant popularity within the computational mechanics community for modeling fractures. It effectively simulates crack initiation, propagation, branching, and merging, eliminating the need for explicit ad-hoc criteria. While this approach eliminates the complexity of constantly modifying the mesh as cracks develop, it introduces the challenge of dealing with the highly nonlinear, non-convex, and non-smooth characteristics of the underlying energy function. To tackle this optimization challenge, we employ a field-split-based additive/multiplicative Schwarz preconditioned Newton method. We will validate the robustness and efficiency of our proposed method by benchmarking it against the conventional alternate minimization approach.

Title: Accelerating training of physics-informed neural networks using decomposition strategies

Alena Kopanicakova

Università della Svizzera italiana

Abstract:

In this talk, we will discuss nonlinear preconditioner-based training for physics-informed neural networks (PINNs). Here, we introduce nonlinear additive and multiplicative preconditioning strategies tailored for the popular L-BFGS optimizer. These preconditioners are built using the Schwarz domain-decomposition framework, allowing for a layer-wise decomposition of the network's parameters. Our numerical experiments show that both additive and multiplicative preconditioners significantly improve the convergence rates over the standard L-BFGS optimizer. These preconditioners not only enhance training speed but also improve accuracy by providing more precise solutions to the underlying partial differential equations. This is joint work with Hardik Kothari, George Karniadakis and Rolf Krause.

Title: Adaptive optimised Schwarz methods

Conor McCoid

Université Laval

Abstract:

Optimized Schwarz methods use Fourier analysis or similar to find transmission conditions between subdomains that provide faster convergence over standard Schwarz methods. However, this requires significant upfront analysis of the operator, and may not be straightforward for all problems. This work presents black box methods for adaptively optimizing the transmission conditions, which is equivalent to a Krylov subspace method.

MS 12. Preconditioning Techniques for Gaussian Processes

Title: On Gaussian Kernel Matrices: Spectral Properties and Efficient Approximations

Difeng Cai

Southern Methodist University

Abstract:

Matrices associated with exponential-type kernels such as Gaussians are commonly found in physics, uncertainty quantification and machine learning. Understanding the spectral and approximation properties of such matrices is important for the design of efficient algorithms and preconditioning techniques. In this talk, we first discuss the spectral properties of kernel matrices with Gaussian or exponential kernels in relation to the corresponding integral operator. Then we show how the matrix structure changes with hyperparameters in the kernel and approximation scheme. Applications of theoretical results to preconditioning will also be presented.

Title: Spectral Shape Estimation of Kernel Matrices

Mikhail Lepilov

Emory University

Abstract:

Kernel matrices of data sampled from some latent distribution appear frequently in data science, but these are often too large to even store in memory, let alone perform computations with. We would, therefore, like to know in advance if it is possible to find an accurate low-rank representation of the matrix in question, or otherwise if doing computations with the matrix is infeasible. That is, we would like to know if a given kernel matrix has low numerical rank. In this work, we explore probabilistic ways to approximate the whole spectrum of a kernel matrix given access to the distribution that the matrix comes from and the kernel to be used. In doing so, we propose a new quantile-based framework to measuring and ensuring the closeness of an eigenvalue distribution of a matrix, as well as some applications thereof.

Title: Efficient Preconditioned Unbiased Estimators in Gaussian Processes

Tianshi Xu

Emory University

Abstract:

Hyperparameter tuning is crucial in Gaussian Process (GP) modeling for achieving accurate predictions. Existing methods often face a trade-off between bias and variance, with traditional approaches introducing bias and randomized-truncated CG (RT-CG) suffering from high variance. In this talk, we introduce the Preconditioned Single-Sample CG (PredSS-CG) estimator, designed to reduce variance while maintaining unbiasedness, thus allowing GP models to handle more complex datasets. We demonstrate the effectiveness of PredSS-CG in

accurately estimating Log Marginal Likelihood (LML) and its gradient on several real-world datasets. This research was collaboratively executed with Hua Huang, Shifan Zhao, Edmond Chow, and Yuanzhe Xi.

MS 13. Recent Progress on Learning to Precondition with Graph Neural Networks

Title: Graph neural network based preconditioner for Krylov subspace methods

Paul Häusner

Uppsala University

Abstract:

Graph neural networks (GNNs) are one of the most popular neural network architectures emerging in the last couple of years. This is owed in part to their adeptness at handling unstructured inputs, a common feature in many real-world scenarios. Moreover, given that many classical algorithms can be framed within the realm of graph problems, GNNs emerge as a natural option for accelerating or substituting traditional algorithms with neural network approaches. In this talk, we first discuss the strong connections between problems arising in numerical linear algebra and the message-passing scheme implemented by many modern GNN models. Then, we showcase how this connection can be exploited in order to efficiently learn preconditioners for Krylov subspace methods using graph neural networks as a computational backend. By choosing a problem-specific architecture and efficient to compute loss, we train a model to predict the incomplete factorization of an input matrix for problems arising from a problem distribution. During inference, we are then able to produce effective preconditioners for unseen problems with a small computational overhead. This allows us to accelerate the total solving times of linear equation systems compared to employing classical general-purpose preconditioning techniques.

Title: Graph Neural Networks for Selection of Preconditioners and Krylov Solvers

Ziyuan Tang

University of Minnesota

Abstract:

Solving large sparse linear systems is a common task in science and engineering, generally necessitating the use of iterative solvers and preconditioners due to the inefficiency of using direct solvers. The practical performance of these solvers and preconditioners is often beyond theoretical analysis, requiring intuition from domain experts, knowledge of hardware, and extensive trial and error. In this work, we introduce a novel method for automatically selecting solver-preconditioner pairs using graph neural networks (GNNs) as a complementary solution to laborious expert efforts. This method begins by unifying sparse matrices of varying sizes into a consistent graph representation with a set of predefined node and graph features. By leveraging the graph structure through a message-passing mechanism, node and graph features are

integrated via graph convolutions. A two-level pooling is also introduced as an extension to standard GNNs. The output embeddings can then be effectively used for classification tasks. Numerical results show that the proposed model is comparable to traditional machine learning models in the Label Ranking Average Precision (LRAP) evaluation metric and outperforms them in the Normalized Discounted Cumulative Gain (NDCG) evaluation metric.

Title: Approximating the Inverse of a Sparse Linear Operator with Graph Neural Networks

Jie Chen

IBM Research

Abstract:

Preconditioning is at the heart of the iterative solutions of large, sparse linear systems of equations. We consider general-purpose preconditioners applicable to many applications. In this case, the assumed knowledge is only the matrix (and the right-hand side) but not the domain or application. We study the use of graph neural networks (GNNs) as an approximation of the matrix inverse, because a graph is naturally associated with the matrix, just like in algebraic multigrid. We build GNNs, propose training methods, and investigate how GNNs behave by vetting a significant portion of the SuiteSparse matrix collection (nearly a thousand matrices). We conclude that GNNs are useful for solving challenging problems and suggest future directions for the research.

Nearby Restaurants

Lunch options:

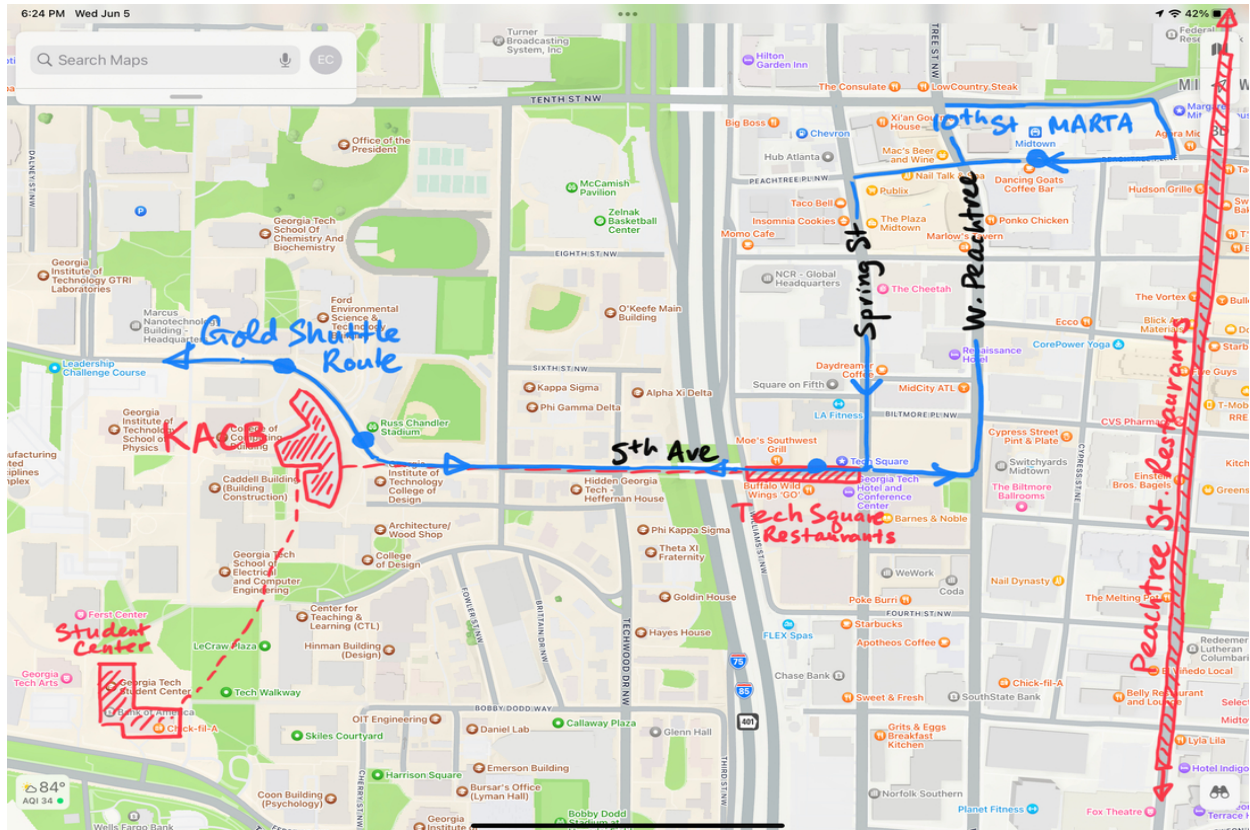
- **The student center** is 8 minutes away and has a variety of student food options on the first and second floors. From KACB, walk towards Tech Green and the student center is the building behind the Campanile Fountain. Note that some dining options are closed during the summer:

<https://dining.gatech.edu/dining-locations/hours-operation#john-lewis-student-center>

- **Tech Square** is 10 minutes away and also has many dining options, mostly on 5th St between Williams St and Spring St.

Dinner options:

Midtown Atlanta offers many fantastic dining options. The main stretch is along Peachtree St NE (not to be confused with West Peachtree St NE runs parallel two blocks west) between North Avenue and 15th Street. Note that many restaurants are closed Mondays.



Directions to the Conference Dinner

[South City Kitchen Midtown, 1144 Crescent Ave NE:](#)



From KACB, South City Kitchen is a 30 minute walk. The map here shows a suggested scenic and pleasant route. It is also possible to take the free Georgia Tech Gold shuttle which goes from KACB to the Midtown MARTA station (blue dot on map) which is the final stop. From there, the walk to the restaurant is much shorter.

<http://bus.gatech.edu/routes/9/stops>