Sparse days meeting 2025

2-3 June @ CERFACS

CERFACS

IRIT

Abstracts

SESSION 1 — Combinatorics and basic sparse matrix manipulation

Sparse matrix computations for verifying the binary Goldbach conjecture

Rob BISSELING

Mots-clés : Goldbach conjecture, number theory, parallel computing, sparse matrix, hypergraph vertex cover problem

Résumé :
The binary Goldbach conjecture states that every even number larger than two can be written as the sum of two prime numbers. The conjecture is an open problem in mathematics since 1742. Its truth has been verified for even numbers up to 4 x 1018 in April 2012, but a proof has not yet been found. We will present a new algorithm for verifying the conjecture, which is based on the use of a suitable sparse data structure, exploitation of repeating patterns in the prime numbers, their translation into a sparse matrix, and several techniques from the area of combinatorial scientific computing, in particular for solving the hypergraph vertex cover problem. Furthermore, we harness parallel computers to reach the high numbers we are interested in. Let\\\'s see how far we can get!

Sparse Matrix Representations Using Minimal Perfect Hashing

Sandeep KORANNE

Mots-clés : operator representations, minimal perfect hashing, Krylov space methods, HPC, indirect memory addressing bottlenecks

Résumé :
Finite difference discretizations of 3D problems create large sparse matrices with few unique non-zero (NZ) entries. Usually, these matrices are stored using compressed row/column industry standard formats such as ELLPACK, YALE, or dictionary-of-keys (DOK). The indirect addressing of these formats does not perform well on modern hierarchical memory processors. In this paper we present an optimization based on minimal perfect hashing to store sparse matrices, especially those derived from finite difference stencils. Our implementation provides a substantial speedup over conventional methods. Examples from Krylov methods and sparse linear solvers are presented. Similar sparse matrices are also used in computational chemistry and we believe our method can be used in that context as well. In this paper an alternative representation using minimal perfect hashing is developed which provides advantages on hierarchical memory processors (most modern CPUs have orders of magnitude latency difference between L1, L2 cache vs main memory) by trading CPU instructions for indirect addressing modes used by traditional SpM storage schemes. During the construction of the SpM hash we also generate an unique key which encodes the hash-function, g array as well as the ROW, COL and BITSET information. This KEY is now part of the C++ type and the C++ language will prevent any misuse of operator+() across non-conforming

minSpM. The advantage of this restriction is that almost all element-wise operations can be done on the NZ store directly. All operations such as +,-,,/,,| can be done simply. Numerical results using C++ and comparing with Eigen library are shown N L2 DCM Ratio NNZ Eigen3 minSpM 16384 0.9 63884 3.27 1.99 65536 0.8 255850 46.9 30.6 262144 0.7 1024301 820 583

A composable abstraction of hierarchical methods for matrix-vector product acceleration
Antoine GICQUEL

Mots-clés : Matrix-vector product, Data-sparse Matrix, Fast Multipole Method, Hierarchical Matrices, Low-rank approximation, Composability

Résumé :
The matrix-vector multiplication is a fundamental mathematical operation in various fields of scientific computing, acting as the cornerstone of numerous iterative algorithms. Thus, accelerating this operation can significantly enhance the computational efficiency of multiple solvers. When working with data-sparse matrices characterized by block low-rank structures, hierarchical methods are often employed as they reduce the computational cost and memory requirements from quadratic to log-linear or even linear complexity with high accuracy. Some common examples include the Fast Multipole Method (FMM), \\\\mathcal{H} matrices, \\\\mathcal{H}2 matrices, Hierarchically Semiseparable (HSS) matrices, Hierarchically Off-Diagonal Low-Rank (HODLR) matrices, and Block Low-Rank (BLR) matrices, which are sometimes referred to as their flat counterparts. All of these hierarchical methods rely on similar core components they partition the matrix into blocks and leverage low-rank approximations of off-diagonal blocks that can be combined with a nested-basis strategy. Despite these similarities, each method defines its own matrix-vector product algorithm. In this work, we propose a composable abstraction that encompasses and generalizes all of these hierarchical methods. From this abstraction, we derive a single generic matrix-vector product algorithm that is modular, with each component being specialized to address different scenarios. Experimental validation demonstrates the efficacy of our abstraction in achieving comparable numerical results to the existing original algorithms.

Recent Results on Improving Performance of Sparse Cholesky Factorization by Reordering Columns within Supernodes
Esmond NG

Mots-clés : right-looking sparse Cholesky with blocking, supernodes, reordering within supernodes, partition refinement, traveling salesman problem

Résumé :
In some recent papers, researchers have found two very good methods for reordering columns within supernodes in sparse Cholesky factors; these reorderings can be very useful for certain factorization methods. The first of these reordering methods is based on modeling the underlying problem as a traveling salesman problem (TSP), and the second of these methods is based on partition refinement (PR). We compare the two methods in this talk using an implementation of sparse Cholesky factorization whose performance depends on reordering columns within supernodes. While the two methods are virtually the same in the quality of the reorderings that they produce, PR should be the method of choice because PR reorderings can be computed using far less time and storage than TSP reorderings.

SESSION 2 — Sparse direct methods

Sherry LI	Batched Sparse Direct Solver Design and Evaluation in SuperLU DIST
Antoine JEGO	BLAS-based Block Memory Accessors with Applications to Mixed-Precision Sparse Direct Solvers
George DELIC	Accelerating Sparse Matrix Algorithms in Global Climate and Air Quality Models
Weifeng LIU	Scaling Up Sparse Direct Solvers on GPU Clusters

Batched Sparse Direct Solver Design and Evaluation in SuperLU DIST
Sherry LI

Mots-clés :

Résumé :
Over the course of interactions with various application teams, the need for batched sparse linear algebra functions has emerged in order to make more efficient use of the GPUs for many small and sparse linear algebra problems. In this talk, we present our recent work on a batched sparse direct solver for GPUs. The sparse LU factorization is computed by the levels of the elimination tree, leveraging the batched dense operations at each level and a new batched Scatter GPU kernel. The sparse triangular solve is computed by the level sets of the directed acyclic graph (DAG) of the triangular matrix. Batched operations overcome the large overhead associated with launching many small kernels. For medium sized matrix batches with not-so-small bandwidth, using an NVIDIA A100 GPU, our new batched sparse direct solver is orders of magnitude faster than a batched banded solver and uses less than one-tenth of the memory.

BLAS-based Block Memory Accessors with Applications to Mixed-Precision Sparse Direct Solvers
Antoine JEGO

Mots-clés : Mixed Precision, BLAS, Block Low Rank, Sparse solver

Résumé :
Mixed-precision algorithms can take advantage of the lower precision of operands to enhance performance. In various contexts, it is beneficial to decouple the storage precision from the compute precision the data is stored and accessed in low precision, but the computations are kept in high precision. This ``memory accessor'' approach benefits from reduced data accesses and improved accuracy, and can simplify the programming of mixed precision software packages. In this work, we develop such a memory accessor and investigate how it can accelerate sparse linear solvers. In particular, we assess its impact on custom floating-point datatypes unsupported by hardware and on structures such as Block Low-Rank formats. When considering BLAS-2 memory-bound operations like trsv we observe that the storage cost adequately matches the performance of the operation, in multiple parallel settings, provided that the conversion from storage to compute precision is efficient. For custom datatypes, we may take advantage of the recent AVX512-VBMI instruction set to reach the required efficiency. We present performance experiments using the sparse solver MUMPS.

Accelerating Sparse Matrix Algorithms in Global Climate and Air Quality Models
George DELIC

Mots-clés : GPU, Sparse Matrix, High Performance Computing

Résumé :
Global Climate Models and Air Quality Models require solution of aqueous chemistry as a system of nonlinear ordinary differential equations integrated in time. The system of equations may number in the hundreds and the number of species is also typically counted in the hundreds. The solution is performed in a chemistry transport model (CTM) as a finite rank sparse matrix system that is integrated in time steps within a transport model. This accounts for a large fraction of the total wall-clock time of a simulation and therefore it is of interest if the CTM algorithm may be accelerated when ported to Graphical Processing Units (GPU). This presentation discusses exploration of GPU offloading the compute intensive calculations for the sparse matrix LU decomposition and solution. The GPU results are compared with those of the host CPU for the case of the Gear ODE solver algorithm.

Scaling Up Sparse Direct Solvers on GPU Clusters
Weifeng LIU

Mots-clés : Sparse direct solver, Scale-up, GPU

Résumé :

Existing sparse direct solvers have not been able to well exploit the high bandwidth and floating-point performance of modern GPUs. The primary challenges arise from inherent inefficiencies in processing small tasks. In this presentation, we will introduce how we significantly enhance the execution efficiency of sparse direct solvers on GPU clusters through aggregating and batching of small tasks, and demonstrate how this approach is integrated into existing libraries.

SESSION 3 — Multigrid methods

Petr Vacek	Mixed precision multigrid V-cycle method
Sai Aakash DASARI	Extrapolated Geometric Multigrid Solver for Poisson-like Equations on Tokamak Geometries
Samuel DUARTE DOS SANTOS	Learning Convolutional Smoothing Operators for Multigrid Solvers of the 2D Poisson Equation
Jimmie ADRIAZOLA	Computer Assisted Discovery of Integrability via SILO Sparse Identification of Lax Operators

Mixed precision multigrid V-cycle method

Petr Vacek

Mots-clés : multigrid, mixed precision computation

Résumé :
Modern parallel computers support computing in multiple precisions. There is extensive ongoing research on numerical methods exploiting this hardware feature. Some mixed precision methods are able to achieve the same overall accuracy as their uniform precision counterparts, while being faster, requiring less memory and consuming less energy. In this talk we discuss several ways of utilizing mixed precision in multigrid methods for solving systems of linear equations. We consider the application of multigrid both as a standalone solver or as a preconditioner for an iterative method. We present novel theoretical results and numerical experiments implemented in the Ginkgo library on GPUs.

Extrapolated Geometric Multigrid Solver for Poisson-like Equations on Tokamak Geometries

Sai Aakash DASARI

Mots-clés : Multigrid, Extrapolation, Isoparametric finite elements

Résumé :
Geometric multigrid methods require problem-specific customisation to achieve optimal algebraic convergence on anisotropic grids. Additionally, extrapolation techniques can be integrated into the multigrid framework to improve the differential convergence of discrete solutions. In this presentation, we discuss our ongoing work to develop an extrapolated geometric multigrid solver for Poisson-like equations with sharp contrasts in diffusion coefficients. We are specifically interested in systems arising from isoparametric finite element discretisations, where curved elements are used to approximate complex geometries.

Learning Convolutional Smoothing Operators for Multigrid Solvers of the 2D Poisson Equation

Samuel DUARTE DOS SANTOS

Mots-clés : Multigrid methods, Poisson equation, Deep learning, Relaxation operator, Convolutional smoother stencil, U-Net analogy, PDE solver

Résumé :
Solving elliptic partial differential equations such as the 2D Poisson equation is a fundamental task in computational physics and engineering. Among the most effective solvers, multigrid methods stand out for their ability to handle errors across multiple scales. However, their efficiency heavily relies on the choice of the smoothing operator, which is traditionally based on simple fixed-point schemes like Jacobi or Gauss-Seidel. This work aims to investigate the multigrid performance by replacing the standard relaxation step with a learned convolutional damping operator. In the simplest case (1×1 kernel), this amounts to learning an optimal scalar relaxation factor ω per level of the V-cycle. For larger stencils (e.g., 3×3), the method learns a full convolutional kernel, allowing more expressive local corrections. We train these convolutional kernels using supervised learning on synthetic data. We implement the framework in PyTorch Lightning. Each level of the multigrid hierarchy receives its own learned kernel. We evaluate the convergence speed, analyze the effect of stencil size and number of V-cycles, and compare against standard multigrid.

Computer Assisted Discovery of Integrability via SILO Sparse Identification of Lax Operators

Jimmie ADRIAZOLA

Mots-clés : Hamiltonian Mechanics, Integrability, Sparse Regression

Résumé :
Integrability in Hamiltonian dynamical systems is a mathematically rich topic and often the starting point for analyzing more complex, nonintegrable physical scenarios. However, it is difficult to even recognize if a given system is integrable before investing effort into studying it. Therefore, we formulate the automated discovery of integrability in dynamical systems, specifically as a sparse symbolic regression problem. Loosely speaking, we seek to maximize the compatibility between the known Hamiltonian of the system and a pair of matrix/differential operators known as Lax pairs. Our approach is tested on a variety of systems ranging from nonlinear oscillators to canonical Hamiltonian PDEs. We test robustness of the framework against nonintegrable perturbations, and, in all examples, reliably confirm or deny integrability. Moreover, because we use a thresholded regularization to promote sparsity, the Lax pairs we discover are interpretable which leads to mathematical theorems. We will discuss future directions for adapting our framework toward further automated discoveries in mathematical physics.

SESSION 4 — Approximate computing

Theo MARY	Mixed precision strategies for preconditioned GMRES
Matthieu ROBEYNS	Mixed precision randomized low-rank approximation with GPU tensor cores
Karmijn HOOGVELD	A survey on randomized low-rank approximation algorithms
Tom CARUSO	Mixed Precision in Domain Decomposition Preconditioner Perturbation Analysis and Practical Guidelines
Miroslav TŮMA	Solving sparse linear least-squares using low precision incomplete Cholesky preconditioning

Mixed precision strategies for preconditioned GMRES
Theo MARY

Mots-clés :

Résumé :
Many different strategies have been proposed to exploit mixed precision arithmetic in GMRES. In this talk, we will describe a unifying framework that covers left, right, and flexible preconditioning and that considers six independent precision parameters. In this framework, we carry out an error analysis that reveals the role of each precision on the attainable accuracy of GMRES and leads to the sharpest bounds to date. These bounds prove that each precision behaves very differently, depending on the condition numbers of the matrix, the preconditioner, and the preconditioned matrix, and depending on the preconditioning style (left, right, flexible). Our framework therefore identifies meaningful mixed precision strategies, recovering known ones and unveiling new ones. We show experimentally that these new strategies offer varying levels of tradeoff between performance and accuracy.

Mixed precision randomized low-rank approximation with GPU tensor cores
Matthieu ROBEYNS

Mots-clés : mixed precision algorithms, randomized algorithms, low-rank approximations, GPU computing, tensor cores

Résumé :
Randomized projection methods have been shown to be very efficient at computing low-rank approximations (LRA) of large matrices. In this work, we investigate the design and development of such methods capable of exploiting recent mixed precision accelerators like GPUs equipped with tensor core units. We combine three new ideas to exploit mixed precision arithmetic in randomized LRA. The first is to perform the matrix multiplication with mixed precision fp16/fp32 tensor cores. The second is to use CholeskyQR orthonormalization, which is much faster on GPUs, while mitigating its numerical instability by using fp64 arithmetic. The third is to use a recently proposed iterative refinement method for LRA to improve the accuracy of the LRA by calling it twice. We implement the proposed approach on various GPU architectures and analyze its performance and accuracy. We compare with a standard randomized LRA entirely in fp32 arithmetic, which achieves

an average accuracy of order 10−4. Our results show that our approach without refinement is up to 8× faster, with an average accuracy of order 10−2, which may be acceptable for some applications. Otherwise, we show that using refinement significantly improves the accuracy to an average of order 10−5, while remaining up to 2.2× faster than the standard fp32 randomized LRA. This work illustrates the convergence of approximate computing techniques by combining low-rank approximations, randomization, mixed precision arithmetic, and GPU acceleration.

A survey on randomized low-rank approximation algorithms

Karmijn HOOGVELD

Mots-clés : Low-rank approximation, Randomization, Operational complexity

Résumé :
In our work we take interest in computing low-rank approximations of dense matrices by means of randomized algorithms. These algorithms have become increasingly popular thanks to their relatively low operational complexity and their high efficiency and scalability on parallel, heterogeneous supercomputers. In this talk we will present a survey of some of the most well known methods and a unified operational complexity analysis. Furthermore we will introduce a novel variant of the randomized interpolative decomposition (RID) algorithm that can compute the fixed-accuracy low-rank approximation thanks to an incremental column selection based on updating QR factorization. We will show some preliminary experimental results and conclude with future research directions.

Mixed Precision in Domain Decomposition Preconditioner Perturbation Analysis and Practical Guidelines

Tom CARUSO

Mots-clés : Domain Decomposition, Mixed Precision Arithmetic, Perturbation Theory, Sparse Linear Systems

Résumé :
This work is concerned with domain decomposition preconditioners, specifically the Schwarz method with GENEO coarse space, to solve large, sparse symmetric positive definite problems. Through libraries such as HPDDM, these preconditioners have already been efficiently parallelized in their numerical implementations. However, they still require expensive linear algebra operations in each local subdomain. Motivated by the emergence of fast low precision arithmetic in hardware, we aim in this work to speed them up using mixed precision. To do this, we need to identify the sensitivity of these operations to perturbation and propose an actionable criteria for selecting the appropriate precision for each local subdomain. In order to do so, we develop a perturbation theory for our preconditioner to bound the worst-case loss of efficiency of the preconditioner and study the sharpness of this theoretical bound through numerical experiments with the FreeFEM, petsc4py, and HPDDM libraries. Our findings show that the only important parameter is the maximum of the sizes of the local subdomain perturbations, weighted by the condition number of the local subdomain matrix. Our results therefore suggest that preconditioners can be constructed in mixed precision while effectively controlling the loss of efficiency.

Solving sparse linear least-squares using low precision incomplete Cholesky preconditioning

Miroslav TŮMA

Mots-clés : sparse least squares, preconditioning, mixed precision

Résumé :
Hardware developments have led to a surge in interest in exploiting mixed precision arithmetic within numerical linear algebra algorithms to take advantage of potential savings in memory requirements, runtime and energy use, whilst still achieving the requested accuracy. The talk will discuss employing mixed precision when solving least-squares problems, focusing on the practicalities of developing robust approaches using low precision incomplete Cholesky factorization preconditioners. Key penalties associated with lower precision include a loss of reliability and less accuracy in the computed solution. Through experiments involving problems from practical applications, we study computing incomplete Cholesky factorizations of the normal matrix using low precision and using the factors to precondition LSQR using mixed precision. We investigate level-based and memory-limited incomplete factorization preconditioners. We find that the former are not effective for solving challenging least-squares problems while the latter can provide high-quality preconditioners. In particular, half precision arithmetic can be considered if high accuracy is not required in the solution or the memory for the incomplete factors is very restricted; otherwise, single precision can be used, and double precision accuracy recovered while reducing memory consumption, even for ill-conditioned problems.

SESSION 5 — HPC/GPU

Joost ROMMES	Accelerated sparse linear algebra and related developments for GPU and hybrid CPU+GPU computing
Yu-Hsiang TSAI	Recent Advancements in the Sparse Linear Algebra Library Ginkgo
Stojche NAKOV	Scalable Domain Decomposition Methods for Large Sparse Linear Systems on Modern Architectures

Accelerated sparse linear algebra and related developments for GPU and hybrid CPU+GPU computing

Joost ROMMES

Mots-clés : GPU, hybrid CPU+GPU, cuDSS, accelerated computing, sparse linear algebra, emulation

Résumé :
GPUs and CPU+GPU superchips provide great opportunities for accelerating scientific computing applications in areas like Electronic Design Automation (EDA), Computational Fluid Dynamics (CFD), and Computer-Aided Engineering (CAE). We will show recent developments and results for new chips like Blackwell and Grace-Blackwell for several of NVIDIA's sparse linear algebra libraries. For the direct sparse solver library cuDSS, we will present how through hybrid CPU+GPU computing speed-ups up to 4x can be obtained on Grace-Blackwell and how systems requiring several hundreds of GB of memory can be solved. Furthermore, we will demonstrate energy-efficient supercomputing with floating point emulation, giving up to 3x speedups over native FP32 and FP64 on Blackwell, and up to 2x power-efficiency. We will also highlight how with nvmath-python we provide Python applications with high-performance pythonic access to our (sparse and dense) linear algebra libraries.

Recent Advancements in the Sparse Linear Algebra Library Ginkgo

Yu-Hsiang TSAI

Mots-clés : GPU, linear solver, mixed precision

Résumé :
Ginkgo is a portable and high-performance library designed for sparse linear algebra, with a focus on GPU optimization. It implements several algorithms, including the Conjugate Gradient (CG), Generalized Minimal Residual Method (GMRES), and Algebraic Multigrid (AMG), in highly efficient ways. Additionally, Ginkgo explores innovative techniques to enhance performance, such as mixed precision, batch functionality, and direct solvers. We have recently added support for half and bfloat16 precision types across most functionalities in Ginkgo. Furthermore, we have improved the usability of the Ginkgo Library by allowing configuration from file and adding Python bindings. In this presentation, we will discuss the new features and enhancements introduced in Ginkgo.

Scalable Domain Decomposition Methods for Large Sparse Linear Systems on Modern Architectures

Stojche NAKOV

Mots-clés : multicore; GPU; domain decomopision methods; sparse hybrid solver

Résumé :
In this work, we address the challenge of scaling domain decomposition methods for solving large sparse linear systems on modern architectures. We leverage the Composyx software framework, integrating state-of-the-art libraries such as BLAS/LAPACK for dense linear algebra, direct sparse solvers for subproblem resolution, and efficient partitioners for domain decomposition. Ensuring seamless integration and optimal resource utilization is critical for computational efficiency. We first optimize these methods for homogeneous multi-core systems, refining each algorithmic stage to minimize overhead and fully exploit shared memory within each domain. A key aspect of this optimization is ensuring that each library efficiently utilizes multi-core architectures through Composyx, both independently and within a shared execution context. Additionally, proper thread and process placement is crucial, ensuring that each process is bound to a specific portion of the machine for optimal performance. To better understand the effectiveness of these optimizations, we will conduct an in-depth analysis of the trade-offs between computational and numerical factors, specifically the number of cores assigned per domain versus the number of domains. The goal is to identify the best configuration that balances parallel efficiency, numerical accuracy, and overall solver performance. This study will provide valuable insights into how domain decomposition methods can be adapted to different problem sizes and hardware configurations, ensuring robustness and scalability. Building on this foundation, ongoing work focuses on extending our framework to heterogeneous CPU-GPU architectures through a fine-grained task-based parallelization strategy. By addressing these challenges related to increasing platform heterogeneity and algorithmic complexity, this work seeks to ensure that domain decomposition methods remain robust, scalable, and well-suited to the computational demands of large-scale simulations in the exascale era.

SESSION 6 — GMRES

Andrew WATHEN	Preconditioning and iteration for indefinite linear systems
Yongseok JANG	Combining mixed precision with deflation and augmentation for accelerating GMRES
Alexandre MALHÉNÉ	Practical backward error based stopping criteria for GMRES
Faris ATTALLAH	Accelerating Power Flow Solvers GPBiCG and BiCGSTAB in Newton-Raphson Methods for Large-Scale Systems

Preconditioning and iteration for indefinite linear systems
Andrew WATHEN

Mots-clés : preconditioning, convergence of iterative methods

Résumé :
The solution of systems of linear(ized) equations lies at the heart of many problems in Scientific Computing. In particular for systems of large dimension, iterative methods are an important approach. Stationary iterative methods are generally based on a matrix splitting, whereas for polynomial iterative methods such as Krylov subspace iteration, the splitting matrix is the preconditioner. The smoother in a multigrid method is generally a stationary or polynomial iteration. Here we consider real symmetric indefinite and complex Hermitian indefinite coefficient matrices and prove that no splitting matrix can lead to a contractive stationary iteration unless the inertia is exactly preserved. This has consequences for preconditioning for indefinite systems and smoothing for multigrid as we further describe.

Combining mixed precision with deflation and augmentation for accelerating GMRES
Yongseok JANG

Mots-clés : Mixed-precision, GMRES, deflated restart, augmented Krylov method

Résumé :
We present a new framework for a mixed precision generalized minimal residual (GMRES) algorithm, focusing on the integration of deflated restart (GMRES-DR) and augmentation (AugGMRES) techniques. We first explore the impact of the preconditioning precision and propose a flexible deflated GMRES (FGMRES-DR) with lower precision preconditioning which preserves a fast convergence. We then explore also performing the orthonormalization process in lower precision, and provide an analysis that shows that unfortunately this leads to inaccuracies in the deflation procedure and stagnation in convergence. To overcome this, we turn to augmentation techniques (AugGMRES), which we show to be able to tolerate a lower precision orthonormalization. In our numerical experiments, we compare the performance of AugGMRES and FGMRES-DR across various mixed precision configurations. Our results show that AugGMRES consistently succeeds in combining the faster convergence of augmented methods with the lower per-iteration cost of mixed precision methods. This new mixed precision augmented GMRES algorithm therefore provides a robust and efficient solution for large, sparse linear systems, offering significant benefits for scientific computing and engineering applications.

Living in a Heterogenous World: How scientific workflows bridge diverse cyberinfrastructure and what we can do better?

Alexandre MALHÉNÉ

Mots-clés : GMRES, stopping criteria, QR factorization

Résumé :

The Generalized Minimal Residual (GMRES) iterative solver is a widely used subspace method for solving large linear systems. The backward stability of its Modified Gram-Schmidt and Householder variants has been established based on the Frobenius norm of the matrix involved in the system. However, computing this norm can be challenging, especially when the matrix is not explicitly available, as is often the case in applications such as non-overlapping domain decomposition or fast multipole methods. In this talk, we will explore several alternatives that leverage by-products of the GMRES algorithm to obtain sufficiently accurate estimates of the norm, enabling the design of practical stopping criteria.

Accelerating Power Flow Solvers GPBiCG and BiCGSTAB in Newton-Raphson Methods for Large-Scale Systems

Faris ATTALLAH

Mots-clés : Krylov subspace method, GPBICG, Precondition, Power flow problem

Résumé :

Efficiently solving large-scale nonlinear algebraic equations is essential for real-time analysis and planning in contemporary power systems. The Newton-Raphson (NR) approach is extensively employed for power flow analysis because of its quadratic convergence; nonetheless, each NR iteration necessitates the resolution of a substantial, sparse, and frequently non-symmetric linear system. Conventional direct solvers become computationally prohibitive as the scale of the system increases. This presentation examines the utilization of two sophisticated iterative Krylov subspace methods, Bi-Conjugate Gradient Stabilised (BiCGSTAB) and Generalised Product-type Bi-Conjugate Gradient (GPBiCG), as linear solvers in the NR framework. These methods are ideal for large-scale systems and can utilize sparse matrix structures and preconditioning techniques like ILU(p) to substantially enhance convergence speed. Numerical investigations on benchmark power system models illustrate the efficacy and resilience of these strategies in decreasing computing time while preserving accuracy. The research emphasizes the practical aspects of incorporating iterative solvers into power flow analysis and offers insights into their relative performance.