

RESEARCH CENTRE

**Inria Centre at the University of
Bordeaux**

IN PARTNERSHIP WITH:

**Airbus Central Research & Technology,
Centre Européen de Recherche et de
Formation Avancée en Calcul Scientifique**

2024

ACTIVITY REPORT

Project-Team

CONCACE

**Numerical and Parallel Composability for
High Performance Computing**

DOMAIN

**Networks, Systems and Services,
Distributed Computing**

THEME

**Distributed and High Performance
Computing**

Inria

Contents

Project-Team CONCACE	1
1 Team members, visitors, external collaborators	2
2 Overall objectives	3
3 Research program	5
3.1 Numerical algebra methodologies in model and data-driven scientific computing	5
3.1.1 Scientific computing in large size linear algebra	6
3.1.2 Scientific computing in large dimension multi-linear algebra	6
3.1.3 Scientific continuum between large size and large dimension	7
3.2 Composition of parallel numerical algorithms from a sequential expression	7
4 Application domains	8
4.1 Aeroacoustics Simulation	8
5 Highlights of the year	9
6 New software, platforms, open data	9
6.1 New software	9
6.1.1 composyx	9
6.1.2 ScalFMM	9
6.1.3 CPPDiodon	10
6.1.4 FMR	11
7 New results	11
7.1 Multifacets of lossy compression for scientific data in the Joint-Laboratory of Extreme Scale Computing	11
7.2 Preconditioners based on Voronoi quantizers of random variable coefficients for stochastic elliptic partial differential equations	11
7.3 Robustness and reliability of state-space, frame-based modeling for thermoacoustics	12
7.4 Computing WSBM marginals with Tensor-Train decomposition	12
7.5 A filtered multilevel Monte Carlo method for estimating the expectation of cell-centered discretized random fields	12
7.6 Multilevel Surrogate-based Control Variates	13
7.7 Quadrature Rules in General Continuous Bayesian Networks : Discrete Inference without Discretization	13
7.8 Material Parameter Estimation for a Viscoelastic Stenosis Model Using a Variational Autoencoder Inverse Mapper	14
7.9 Abstracting hierarchical methods for accelerating matrix-vector products on heterogeneous machine clusters: A unified framework for Fast Multipole Methods, \mathcal{H} -Matrices, and beyond	14
7.10 Using the Fast Multipole Method to Accelerate Matrix-Vector Products	14
7.11 Optimizing multigrid efficiency using machine learning techniques	15
7.12 Comparison of multigrid and machine learning-based Poisson solvers	15
7.13 Statistical machine learning techniques to predict the rank in \mathcal{H} -matrix computation	15
7.14 Convergence analysis of overlapping domain decomposition preconditioners for nonlinear problems	15
7.15 Low-rank tensor solver for magnetostatic problems for electric power applications	16
7.16 Intra-node performance	16
8 Partnerships and cooperations	16
8.1 European initiatives	17
8.1.1 H2020 projects	17
8.1.2 Other european programs/initiatives	18
8.2 National initiatives	18

9 Dissemination	21
9.1 Promoting scientific activities	21
9.1.1 Scientific events: organisation	21
9.1.2 Scientific events: selection	21
9.1.3 Journal	21
9.1.4 Invited talks	21
9.1.5 Scientific expertise	22
9.1.6 Research administration	22
9.2 Teaching - Supervision - Juries	22
9.2.1 Teaching	22
9.2.2 Supervision	22
9.2.3 Juries	23
10 Scientific production	23
10.1 Major publications	23
10.2 Publications of the year	24
10.3 Cited publications	26

Project-Team CONCACE

Creation of the Project-Team: 2022 September 01

Keywords

Computer sciences and digital sciences

- A1.1.4. – High performance computing
- A1.1.5. – Exascale
- A1.1.9. – Fault tolerant systems
- A6.2.5. – Numerical Linear Algebra
- A6.2.7. – High performance computing
- A6.3. – Computation-data interaction
- A7.1. – Algorithms
- A8.2. – Optimization
- A8.10. – Computer arithmetic
- A9.2. – Machine learning
- A9.7. – AI algorithmics
- A9.10. – Hybrid approaches for AI

Other research topics and application domains

- B3.3.1. – Earth and subsoil
- B4.2.2. – Fusion
- B5.2.3. – Aviation
- B5.5. – Materials
- B9.5.1. – Computer science
- B9.5.2. – Mathematics
- B9.5.4. – Chemistry
- B9.5.6. – Data science

1 Team members, visitors, external collaborators

Research Scientists

- Luc Giraud [Team leader, INRIA, Senior Researcher]
- Carola Kruse [Team leader, CERFACS, Senior Researcher]
- Guillaume Sylvand [Team leader, AIRBUS, Senior Researcher]
- Emmanuel Agullo [INRIA, Researcher]
- Pierre Benjamin [AIRBUS, Senior Researcher]
- Olivier Coulaud [INRIA, Senior Researcher]
- Sofiane Haddad [AIRBUS, Senior Researcher]
- Paul Mycek [CERFACS, Senior Researcher]

Post-Doctoral Fellows

- Hadrien Godé [CERFACS, Post-Doctoral Fellow, from Feb 2024]
- Marvin Lasserre [INRIA, Post-Doctoral Fellow, until Sep 2024]
- Stojche Nakov [INRIA, Post-Doctoral Fellow, from Sep 2024]
- Yanfei Xiang [INRIA, Post-Doctoral Fellow, until Nov 2024]

PhD Students

- Theo Briquet [INRIA]
- Hugo Dodelin [INRIA, from Oct 2024]
- El Mehdi Ettaouchi [EDF]
- Antoine Gicquel [INRIA]
- Alexandre Malhene [INRIA, from Oct 2024]
- Amine El Mehdi Zekri [TOULOUSE INP]

Technical Staff

- Hugo Dodelin [INRIA, Engineer, from May 2024 until Sep 2024]
- Esragul Korkmaz [INRIA, Engineer, from Oct 2024]
- Gilles Marait [INRIA, Engineer]

Interns and Apprentices

- Guillaume Dindart [INRIA, Intern, from May 2024 until Aug 2024]
- Aurelien Gauthier [INRIA, Intern, from May 2024 until Aug 2024]
- Alexandre Malhene [INRIA, Intern, from Mar 2024 until Aug 2024]

Administrative Assistant

- Flavie Blondel [INRIA]

External Collaborators

- Luciano Drozda [CERFACS, from Aug 2024]
- Jean-Rene Poirier [TOULOUSE INP]

2 Overall objectives

Over the past few decades, there have been innumerable science, engineering and societal breakthroughs enabled by the development of high performance computing (HPC) applications, algorithms and architectures. These powerful tools have enabled researchers to find computationally efficient solutions to some of the most challenging scientific questions and problems in medicine and biology, climate science, nanotechnology, energy, and environment – to name a few – in the field of **model-driven computing**. Meanwhile the advent of network capabilities and IoT, next generation sequencing, ... tend to generate a huge amount of data that deserves to be processed to extract knowledge and possible forecasts. These calculations are often referred to as **data-driven calculations**. These two classes of challenges have a common ground in terms of numerical techniques that lies in the field of linear and multi-linear algebra. They do also share common bottlenecks related to the size of the mathematical objects that we have to represent and work on; those challenges retain a growing attention from the computational science community.

In this context, the purpose of the `concace` project, is **to contribute to the design of novel numerical tools for model-driven and data-driven calculations** arising from challenging academic and industrial applications. The solution of these challenging problems requires a multidisciplinary approach involving applied mathematics, computational and computer sciences. In applied mathematics, it essentially involves **advanced numerical schemes both in terms of numerical techniques and data representation of the mathematical objects** (e.g., compressed data, low-rank tensor [65, 73, 60] low-rank hierarchical matrices [63, 46]). In computational science, it involves large scale parallel heterogeneous computing and **the design of highly composable algorithms**. Through this approach, `concace` intends to contribute to all the steps that go from the design of new robust and accurate numerical schemes to the flexible implementations of the associated algorithms on large computers. To address these research challenges, researchers from Inria, Airbus Central R&T and Cerfacs have decided to combine their skills and research efforts to create the Inria `concace` project team, which will allow them to cover the entire spectrum, from fundamental methodological concerns to full validations on challenging industrial test cases. Such a joint project will enable a real synergy between basic and applied research with complementary benefits to all the partners. The main benefits for each partner are given below:

- Airbus Central R&T
 - Push our specific needs and use-cases towards the academic world to stimulate research in particular directions;
 - Remain at the level of the scientific state of the art, this collaboration allows us to facilitate the feedback by exposing directly our challenges and industrial applications to eventually facilitate the transfer of research in our design tools;
 - The Inria research model will naturally be extended to Airbus, allowing for the multiplication of ambitious, very upstream and long-term research, while at the same time directly applying to the needs expressed by Airbus;
 - Benefit from the very high-level international network of the Inria team (e.g., Univ. of Tennessee Knoxville, Barcelona supercomputing center, Julich supercomputing center, Lawrence Berkeley National Lab, Sandia National Lab, etc.).
- Cerfacs
 - Join forces, in terms of skills and expertise, with Inria and Airbus to make faster and more effective progress on the research areas addressed by the team;
 - Bring scientific challenges from industrial applications through our privileged relationship with our industrial partners;

- Reciprocally, promote the developed methodologies and the obtained results towards our industrial partners;
 - Naturally interact with the national and european HPC ecosystems, as a member of the EuroHPC national competence center on HPC, to promote the research activities and tools of the team and to meet novel scientific challenges where our methodologies or tools apply.
- Inria
 - Reinforce the impact of our research through a direct contact and close interactions with real scientific and technical challenges;
 - Feed the virtuous feedback cycle between academic research and industrially-relevant applications enabling the emergence of new research avenues;
 - Create a privileged space for an open scientific dialogue enabling the fostering of existing synergies and to create new ones, in particular when one of the industrial partners is a large group whose spectrum of scientific problems is very broad.

In addition to the members of these entities, two other external collaborators will be strongly associated: Jean-René Poirier, from Laplace Laboratory at University of Toulouse) and Oguz Kaya, from LISN (Laboratoire Interdisciplinaire des Sciences du Numérique) at University of Saclay.

The scientific objectives described in Section 4 contain two main topics which cover numerical and computational methodologies. Each of the topic is composed of a methodological component and its validation counterpart to fully assess the relevance, robustness and effectiveness of the proposed solutions. First, we address numerical linear and multilinear algebra methodologies for model- and data-driven scientific computing. Second, because there is no universal single solution but rather a large panel of alternatives combining many of the various building boxes, we also consider research activities in the field of composition of parallel algorithms and data distributions to ease the investigation of this combinatorial problem toward the best algorithm for the targeted problem.

To illustrate on a single but representative example of **model-driven problems** that the joint team will address we can mention one encountered at Airbus that is related to large aero-acoustic calculations. The reduction of noise produced by aircraft during take-off and landing has a **direct societal and environmental impact on the populations (including citizen health)** located around airports. To comply with new noise regulation rules, novel developments must be undertaken to preserve the competitiveness of the European aerospace industry. In order to design and optimize new absorbing materials for acoustics and reduce the perceived sound, one must be able to simulate the propagation of an acoustic wave in an aerodynamic flow: The physical phenomenon at stake is **aero-acoustics**. The complex and chaotic nature of fluid mechanics requires simplifications in the models used. Today, we consider the flow as non-uniform only in a small part of the space (in the jet flow of the reactors mainly) which will be meshed in volume finite elements, and everywhere else the flow will be considered as uniform, and the acoustic propagation will be treated with surface finite elements. This brings us back to the solution of a linear system with dense and sparse parts, an atypical form for which there is no "classical" solver available. We therefore have to work on the coupling of methods (direct or iterative, dense or sparse, compressed or not, etc.), and to **compose different algorithms** in order to be able to handle very large industrial cases. While there are effective techniques to solve each part independently from one another, there is no canonical, efficient solution for the coupled problem, which has been much less studied by the community. Among the possible improvements to tackle such a problem, **hybridizing simulation and learning** represents an alternative which allows one to reduce the complexity by avoiding as much as possible local refinements and therefore reduce the size of the problem.

Regarding **data-driven calculation**, climate data analysis is one of the application domains that generate huge amounts of data, either in the form of measurements or computation results. The ongoing effort between the climate modeling and weather forecasting community to mutualize digital environment, including codes and models, leads the climate community to use finer models and discretization generating an ever growing amount of data. The analysis of these data, mainly based on classical numerical tools with a strong involvement of linear algebra ingredients, is facing new scalability challenges due to this growing amount of data. Computed and measured data have **intrinsic structures** that could be naturally exploited by **low rank tensor representations** to best reveal the hidden structure of the data

while addressing the scalability problem. The close link with the **CECI team at Cerfacs** will provide us with the opportunity to study novel numerical methodologies based on tensor calculation. Contributing to a better understanding of the mechanisms governing the climate change would obviously have significant **societal and economical impacts on the population**. This is just an illustration of a possible usage of our work, we could also have possibly mentioned an on-going collaboration where our tools will be used in the context of a steel company to reduce the data volume generated by IoT to be transferred on the cloud for the analysis. The methodological part described in Section 4 covers mostly two complementary topics: the first in the field of numerical scientific computing and the second in the core of computational sciences.

To sum-up, for each of the methodological contributions, we aim to find at least one dimensioning application, preferably from a societal challenge, which will allow us to validate these methods and their implementations at full-scale. The search for these applications will initially be carried out among those available at Airbus or Cerfacs, but the option of seeking them through collaborations outside the project will remain open. The ambition remains to develop generic tools whose implementations will be made accessible via their deposit in the public domain.

3 Research program

The methodological component of our proposal concerns the expertise for the design as well as the efficient and scalable implementation of highly parallel numerical algorithms. We intend to go from numerical methodology studies to design novel numerical schemes up to the full assessment at scale in real case academic and industrial applications thanks to advanced HPC implementations.

Our view of the research activity to be developed in **CONCACE** is to systematically assess the methodological and theoretical developments in real scale calculations mostly through applications under investigations by the industrial partners (namely Airbus Central R&T and Cerfacs).

We first consider in Section 4.1 topics concerning parallel linear and multi-linear algebra techniques that currently appear as promising approaches to tackle huge problems both in size and in dimension on large numbers of cores. We highlight the linear problems (linear systems or eigenproblems) because they are in many large scale applications the main bottleneck and the most computational intensive numerical kernels. The second research axis, presented in Section 4.2, is related to the challenge faced when advanced parallel numerical toolboxes need to be composed to easily find the best suited solution both from a numerical but also parallel performance point of view.

In short the research activity will rely on two scientific pillars, the first dedicated to the development of new mathematical methods for linear and multilinear algebra (both for model-driven and data-driven calculations). The second pillar will be on parallel computational methods enabling to easily compose in a parallel framework the packages associated with the methods developed as outcome of the first pillar. The mathematical methods from the first pillar can mathematically be composed, the challenge will be to do on large parallel computers thank to the outcome of the second pillar. We will still validate on real applications and at scale (problem and platform) in close collaborations with application experts.

3.1 Numerical algebra methodologies in model and data-driven scientific computing

At the core of many simulations, one has to solve a linear algebra problem that is defined in a vector space and that involves linear operators, vectors and scalars, the unknowns being usually vectors or scalars, *e.g.* for the solution of a linear system or an eigenvalue problem. For many years, in particular in model-driven simulations, the problems have been reformulated in classical matrix formalism possibly unfolding the spaces where the vectors naturally live (typically 3D PDEs) to end up with classical vectors in R^n or C^n . For some problems, defined in higher dimension (*e.g.*, time dependent 3D PDE), the other dimensions are dealt in a problem specific fashion as unfolding those dimensions would lead to too large matrices/vectors. The **concace** research program on numerical methodology intends to address the study of novel numerical algorithms to continue addressing the mainstream approaches relying on **classical matrix formalism** but also to investigate alternatives where the structure of the underlying problem is kept preserved and all dimensions are dealt with equally. This latter research activity mostly

concerns **linear algebra in tensor spaces**. In terms of algorithmic principles, we will lay an emphasis on hierarchy as a unifying principle for the numerical algorithms, the data representation and processing (including the current hierarchy of arithmetic) and the parallel implementation towards scalability.

3.1.1 Scientific computing in large size linear algebra

As an extension of our past and on-going research activities, we will continue our works on numerical linear algebra for model-driven applications that rely on classical vectorial spaces defined on R^n and C^n , where vectors and matrices are classical sparse or dense objects encountered in regular numerical linear algebra computations.

The main numerical algorithms we are interested in are:

- **Matrix decompositions** including classical ones such as the QR factorization that plays a central role in block Krylov solvers [42, 59], randomized range finder algorithms [44, 45], to name a few, as building orthonormal basis of subspaces guarantees numerical robustness. But also other factorizations, not used in classical linear algebra for model-driven calculation, such as non-negative factorization encountered in data-science for multi-variable analysis [58, 52].
- **Iterative solvers** both for linear system solutions and for eigenproblems. Regarding linear systems, we will pay a particular attention to advanced numerical techniques such as multi-level preconditioning, hybrid direct-iterative (both algebraic and PDE driven interface boundary conditions) and the solution of augmented systems (e.g., Karush-Kuhn-Tucker or KKT) [66, 67]. We will investigate variants of nested subspace methods, possibly with subspace augmentation or deflation. In the multiple right-hand sides or left-hand sides cases, we will further study the possible orthogonalization variants and the trade-off between the associated parallel scalability and robustness. A particular attention will be paid to the communication hiding approaches and the investigation of their block extensions. For eigenproblem solutions, we will consider novel nested subspace techniques to further extend the numerical capabilities of the recently proposed AVCI [72, 68] technique as well as contour based integral equations (that intensively use linear systems techniques mentioned above).

In that context, we will consider the benefit of using hybridization between simulation and learning in order to reduce the complexity of classical approaches by diminishing the problem size or improving preconditioning techniques. In a longer term perspective, we will also conduct an active technological watch activity with respect to quantum computing to better understand how such an advanced computing technology can be synergized with classical scientific computing.

3.1.2 Scientific computing in large dimension multi-linear algebra

This work will mostly address linear algebra problems defined in large dimensional spaces as they might appear either in model-driven simulations or data-driven calculations. In particular we will be interested in tensor vectorial spaces where the intrinsic mathematical structures of the objects have to be exploited to design efficient and effective numerical techniques.

The main numerical algorithms we are interested in are:

- Low-rank tensor decompositions for model- and data-driven, some of them rely on some numerical techniques considered in the previous section [54, 57];
- Extension of iterative numerical linear solvers (linear systems and eigensolvers) to tensor vectorial spaces to handle problems that were previously vectorized to be amenable to solution by classical linear algebra techniques;
- Study preconditioning and domain decomposition techniques suited for the solution of stochastic PDEs (encountered in some Uncertainty Quantification context) [77] leading to large dimension or preconditioning based on a low-rank approximation of the tensorization of the dense matrix in Boundary Element Method solver [40, 43, 74].

3.1.3 Scientific continuum between large size and large dimension

Novel techniques for large size and large dimension problems tend to reduce the memory footprint and CPU consumption through data compression such as low-rank approximations (hierarchical matrices for dense and sparse calculation, tensor decomposition [56, 75, 69]) or speed up the algorithm (fast multipole method, randomized algorithm [64, 70] [76, 45]) to reduce the time and energy to solution. Because of the compression, the genuine data are represented with lower accuracy possibly in a hierarchical manner. Understanding the impact of this lower precision data representation through the entire algorithm is an important issue for developing robust, “accurate” and efficient numerical schemes for current and emerging computing platforms from laptop commodity to supercomputers. Mastering the **trade-off between performance and accuracy** will be part of our research agenda [49, 53].

Because the low precision data representation can have diverse origins, this research activity will naturally cover the multi-precision arithmetic calculation in which the data perturbation comes entirely from the data encoding, representation and calculation in IEEE (or more exotic Nvidia GPU or Google TPU) floating point numbers. This will result in variable accuracy calculations. This general framework will also enable us to address soft error detection [39] and study possible mitigation schemes to design resilient algorithms.

3.2 Composition of parallel numerical algorithms from a sequential expression

A major breakthrough for exploiting multicore machine [48] is based on a data format and computational technique originally used in an out-of-core context [62]. This is itself a refinement of a broader class of numerical algorithms – namely, “updating techniques” – that were not originally developed with specific hardware considerations in mind. This historical anecdote perfectly illustrates the need to separate data representation, algorithmic and architectural concerns when developing numerical methodologies. In the recent past, we have contributed to the study of the sequential task flow (STF) programming paradigm, that enabled us to abstract the complexity of the underlying computer architecture [38, 37, 36]. In the concace project, we intend to go further by abstracting the numerical algorithms and their dedicated data structures. We strongly believe that combining these two abstractions will allow us to **easily compose toolbox algorithms and data representations in order to study combinatorial alternatives towards numerical and parallel computational efficiency**. We have demonstrated this potential on domain decomposition methods for solving sparse linear systems arising from the discretisation of PEDs, that has been implemented in the maphys++ parallel package.

Regarding the abstraction of the target architecture in the design of numerical algorithms, the STF paradigm has been shown to significantly reduce the difficulty of programming these complex machines while ensuring high computational efficiency. However, some challenges remain. The first major difficulty is related to the scalability of the model at large scale where handling the full task graph associated with the STF model becomes a severe bottleneck. Another major difficulty is the inability (at a reasonable runtime cost) to efficiently handle fine-grained dynamic parallelism, such as numerical pivoting in the Gaussian elimination where the decision to be made depends on the outcome of the current calculation and cannot be known in advance or described in a task graph. These two challenges are the ones we intend to study first.

With respect to the second ingredient, namely the abstraction of the algorithms and data representation, we will also explore whether we can provide additional separation of concerns beyond that offered by a task-based design. As a seemingly simple example, we will investigate the possibility of abstracting the matrix-vector product, basic kernel at the core of many numerical linear algebra methods, to cover the case of the fast multipole method (FMM, at the core of the ScalFMM library). FMM is mathematically a block matrix-vector product where some of the operations involving the extra-diagonal blocks with hierarchical structure would be compressed analytically. Such a methodological step forward will consequently allow the factorisation of a significant part of codes (so far completely independent because no bridge has been made upstream) including in particular the ones dealing with \mathcal{H} -matrices. The easy composition of these different algorithms will make it possible to explore the combinatorial nature of the possible options in order to best adapt them to the size of the problem to be treated and the characteristics of the target computer. *Offering such a continuum of numerical methods rather than a discrete set of tools is part of the team’s objectives* It is a very demanding effort in terms of HPC software

engineering expertise to coordinate the overall technical effort.

We intend to strengthen our engagement in reproducible and open science. Consequently, we will continue our joint effort to ensure consistent deployment of our parallel software; this will contribute to improve its impact on academic and industrial users. The software engineering challenge is related to the increasing number of software dependencies induced by the desired capability of combining the functionality of different numerical building boxes, e.g., a domain decomposition solver (such as `maphys++`) that requires advanced iterative schemes (such as those provided by `fabulous`) as well as state-of-the-art direct methods (such as `pastix`, `mumps`, or `qr_mumps`), deploying the resulting software stack can become tedious [41].

In that context, we will consider the benefit of using hybridization between simulation and learning in order to reduce the complexity of classical approaches by diminishing the problem size or improving preconditioning techniques. In a longer term perspective, we will also conduct an active technological watch activity with respect to quantum computing to better understand how such a advanced computing technology can be synergized with classical scientific computing.

4 Application domains

Participants: Emmanuel Agullo, Théo Briquet, Olivier Coulaud, Antoine Gicquel, Sofiane Haddad, Carola Kruse, Paul Mycek, Pierre Benjamin, Luc Giraud, Gilles Marait, Guillaume Sylvand, Yanfei Xiang.

We have a major application domain in acoustic simulations that is provided by Airbus CR & T and a few more through collaborations in the context of ongoing projects, that include: plasma simulation (ESA contract and ANR Maturation), Electric device design (ANR TensorVim) and nanoscale simulation platform (ANR Diwina).

4.1 Aeroacoustics Simulation

This domain is in the context of a long term collaboration with Airbus Research Centers. Wave propagation phenomena intervene in many different aspects of systems design at Airbus. They drive the level of acoustic vibrations that mechanical components have to sustain, a level that one may want to diminish for comfort reason (in the case of aircraft passengers, for instance) or for safety reason (to avoid damage in the case of a payload in a rocket fairing at take-off). Numerical simulations of these phenomena plays a central part in the upstream design phase of any such project [50]. Airbus Central R & T has developed over the last decades an in-depth knowledge in the field of Boundary Element Method (BEM) for the simulation of wave propagation in homogeneous media and in frequency domain. To tackle heterogeneous media (such as the jet engine flows, in the case of acoustic simulation), these BEM approaches are coupled with volumic finite elements (FEM). We end up with the need to solve large (several millions unknowns) linear systems of equations composed of a dense part (coming from the BEM domain) and a sparse part (coming from the FEM domain). Various parallel solution techniques are available today, mixing tools created by the academic world (such as the Mumps and Pastix sparse solvers) as well as parallel software tools developed in-house at Airbus (dense solver SPIDO, multipole solver, \mathcal{H} -matrix solver with an open sequential version available online). In the current state of knowledge and technologies, these methods do not permit to tackle the simulation of aeroacoustics problems at the highest acoustic frequencies (between 5 and 20 kHz, upper limits of human audition) while considering the whole complexity of geometries and phenomena involved (higher acoustic frequency implies smaller mesh sizes that lead to larger unknowns number, a number that grows like f^2 for BEM and f^3 for FEM, where f is the studied frequency). The purpose of the study in this domain is to develop advanced solvers able to tackle this kind of mixed dense/sparse linear systems efficiently on parallel architectures.

5 Highlights of the year

- On February 8, Paul Mycek defended his HDR titled "Hierarchical methods for deterministic and stochastic partial differential equations" [21], in front of an international jury (4 foreign members) representing a wide range of scientific fields.
- Concace's research focuses on numerical and parallel composability in scientific algorithms. This work aims at being integrated into the open-source **Composyx** software package, whose version 1.0 was released in 2024. This first release is mainly centered on linear algebra with a strong focus on composability. Its goal is to give users a high-level interface to develop a wide range of algorithms, scaling seamlessly from laptop prototypes to large-scale parallel computations on supercomputers.
- MAMBO ("Advanced Methods for Engine and Aircraft Noise Modeling") is a project funded by the DGAC (Direction Générale de l'aviation civile, french Civil Aviation Authority), bringing together 21 academic and industrial partners from France and Europe, including Airbus, Cerfacs, and Inria. For Concace, the project presents an opportunity to tackle new challenges and develop innovative numerical methods by enhancing the capabilities of its software, particularly in the field of high-performance computing. This activity started in 2024.

6 New software, platforms, open data

6.1 New software

6.1.1 composyx

Name: Numerical and parallel composability for high performance computing

Keywords: Numerical algorithm, Parallel computing, Linear algebra, Task-based algorithm, Dense matrix, Sparse matrix, Hierarchical matrix, FMM, C++

Functional Description: Composable numerical and parallel linear algebra library

URL: <https://gitlab.inria.fr/compose/compose>

Contact: Emmanuel Agullo

6.1.2 ScalFMM

Name: Scalable Fast Multipole Method

Keywords: N-body, Fast multipole method, Parallelism, MPI, OpenMP

Scientific Description: ScalFMM is a software library to simulate N-body interactions using the Fast Multipole Method. The library offers two methods to compute interactions between bodies when the potential decays like $1/r$. The first method is the classical FMM based on spherical harmonic expansions and the second is the Black-Box method which is an independent kernel formulation (introduced by E. Darve @ Stanford). With this method, we can now easily add new non oscillatory kernels in our library. For the classical method, two approaches are used to decrease the complexity of the operators. We consider either matrix formulation that allows us to use BLAS routines or rotation matrix to speed up the M2L operator.

ScalFMM intends to offer all the functionalities needed to perform large parallel simulations while enabling an easy customization of the simulation components: kernels, particles and cells. It works in parallel in a shared/distributed memory model using OpenMP and MPI. The software architecture has been designed with two major objectives: being easy to maintain and easy to understand. There is two main parts: the management of the octree and the parallelization of the method the kernels. This new architecture allow us to easily add new FMM algorithm or kernels and new paradigm of parallelization.

The version 3.0 of the library is a partial rewriting of the version 2.0 in modern C++ (C++17) to increase the genericity of the approach. This version is also the basic framework for studying numerical and parallel composability within Concace.

Functional Description: Compute N-body interactions using the Fast Multipole Method for large number of objects

Release Contributions: ScalFmm is a high performance library for solving n-body problems in astrophysics and electrostatics. It is based on the fast multipole method (FMM) and is highly parallel

News of the Year: Performance improvements in version 3.0. For the moment, this version only considers the interpolation approach. New features - the target particles can be different from the source particles - possibility to consider a non-mutual approach in the direct field - the low rank approximation of the transfer operator is taken into account.

URL: <https://gitlab.inria.fr/solverstack/ScalFMM>

Publications: [hal-00911856](#), [hal-01359458](#), [hal-00746089](#), [hal-01517153](#), [hal-01474556](#), [hal-01517153](#)

Contact: Olivier Coulaud

Participants: Olivier Coulaud, Pierre Est erie

6.1.3 CPPDiodon

Name: Parallel C++ library for Multivariate Data Analysis of large datasets.

Keywords: SVD, PCA

Scientific Description: Diodon provides executables and functions to compute multivariate data Analysis such as: Singular Value Decomposition (SVD), Principal Component Analysis (PCA) and variants (with different pre-treatments), Multidimensional Scaling (MDS), Correspondence Analysis (CoA), Canonical Correlation Analysis (CCA, future work), Multiple Correspondence Analysis (MCoA, future work). All these methods rely on a Singular Value Decomposition (SVD) of a 2D matrix. For small size matrices the SVD can be directly computed using a sequential or multi-threaded LAPACK solver such as OpenBlas or Intel MKL. For large matrices the SVD becomes time consuming and we use a Randomized Singular Value Decomposition method (rSVD) instead of the exact SVD which implementation is given by the FMR library. FMR can perform computations of the rSVD on parallel shared and distributed memory machines using adequate parallel dense linear algebra routines internally such as OpenBlas or Intel MKL on a shared memory node and Chameleon for distributed memory nodes (MPI).

Functional Description: Dimension reduction by multivariate data analysis. Diodon is a list of functions and drivers that implement in C++ and Python (i) pre-processing, SVD and post-processing with a wide variety of methods, (ii) random projection methods for SVD execution which allows to circumvent the time limitation in the calculation of the SVD, and (iii) a C++ implementation of the SVD with random projection to an imposed range or precision, connected to the MDS, PCA, CoA.

Release Contributions: Initial release of cppdiodon : a parallel C++ library for Multivariate Data Analysis of large datasets. Contains methods to compute Singular Value Decomposition (SVD), Randomized SVD, Principal Component Analysis (PCA), Multidimensional Scaling (MDS) and Correspondence Analysis (CoA). Handles text and hdf5 files. Parallel (mpi, threads, cuda) randomized SVD and EVD (for symmetric matrices) provided by FMR. Use multithreaded Lapack or Chameleon (distributed systems + GPUs).

URL: <https://gitlab.inria.fr/diodon/cppdiodon>

Publication: [hal-03773985](#)

Contact: Olivier Coulaud

Partner: INRAE

6.1.4 FMR

Name: Fast Methods for Randomized numerical linear algebra

Keyword: SVD

Scientific Description: Fast Dense Standard and Randomized Numerical Linear Algebra is a library that allows to compute singular values or eigenvalues of large dense matrices by random linear algebra techniques. It is based on the random projection method (Gaussian or fast Hadamard/Fourier) or row/column selection (Nystrom method and variants). The library is developed in C++ and proposes a shared memory parallelization and a distributed approach with Chameleon (<https://gitlab.inria.fr/solverstack/chameleon>).

Functional Description: Fast Dense Standard and Randomized Numerical Linear Algebra is a library that allows to compute singular values or eigenvalues of large dense matrices by random linear algebra techniques. It is based on the random projection method (Gaussian or fast Hadamard/Fourier) or row/column selection (Nystrom method and variants). The library is developed in C++ and proposes a shared memory parallelization and a distributed approach with Chameleon (<https://gitlab.inria.fr/solverstack/chameleon>).

URL: <https://gitlab.inria.fr/compose/legacystack/fmr>

Publications: [hal-01228519](https://hal.archives-ouvertes.fr/hal-01228519), [hal-01334747](https://hal.archives-ouvertes.fr/hal-01334747), [hal-03773985](https://hal.archives-ouvertes.fr/hal-03773985)

Contact: Olivier Coulaud

Participants: Olivier Coulaud, Florent Pruvost, Romain Peressoni

7 New results

Participants: All team members.

7.1 Multifacets of lossy compression for scientific data in the Joint-Laboratory of Extreme Scale Computing

The Joint Laboratory on Extreme-Scale Computing (JLESC) was initiated at the same time lossy compression for scientific data became an important topic for the scientific communities. The teams involved in the JLESC played and are still playing an important role in developing the research, techniques, methods, and technologies making lossy compression for scientific data a key tool for scientists and engineers. In this paper, we present the evolution of lossy compression for scientific data from 2015, describing the situation before the JLESC started, the evolution of this discipline in the past 8 years (until 2023) through the prism of the JLESC collaborations on this topic and some of the remaining open research questions

For more details on this work we refer to [16].

7.2 Preconditioners based on Voronoi quantizers of random variable coefficients for stochastic elliptic partial differential equations

A preconditioning strategy is proposed for the iterative solve of large numbers of linear systems with variable matrix and right-hand side which arise during the computation of solution statistics of stochastic elliptic partial differential equations with random variable coefficients sampled by Monte Carlo. Building on the assumption that a truncated Karhunen-Loève expansion of a known transform of the random variable coefficient is known, we introduce a compact representation of the random coefficient in the form of a Voronoi quantizer. The number of Voronoi cells, each of which is represented by a centroidal variable coefficient, is set to the prescribed number P of preconditioners. Upon sampling the random variable coefficient, the linear system assembled with a given realization of the coefficient is solved with

the preconditioner whose centroidal variable coefficient is the closest to the realization. We consider different ways to define and obtain the centroidal variable coefficients, and we investigate the properties of the induced preconditioning strategies in terms of average number of solver iterations for sequential simulations, and of load balancing for parallel simulations. Another approach, which is based on deterministic grids on the system of stochastic coordinates of the truncated representation of the random variable coefficient, is proposed with a stochastic dimension which increases with the number P of preconditioners. This approach allows to bypass the need for preliminary computations in order to determine the optimal stochastic dimension of the truncated approximation of the random variable coefficient for a given number of preconditioners.

For more details on this work we refer to [31].

7.3 Robustness and reliability of state-space, frame-based modeling for thermoacoustics

The Galerkin modal expansion is a well-known method used to develop reduced order models for thermoacoustics. A known issue is the appearance of Gibbs-type oscillations on velocity fluctuations at the interface between subdomains and at boundary conditions. Recent work of Laurent et al. (2019) and Laurent et al. (2021) have shown that it is possible to overcome this issue by using an over-completed frame, instead of a Galerkin modal basis. However, the low-order modeling based on this frame modal expansion may generate spurious modes, that are spurious eigenpairs. In this paper, the origin of these non-physical modes is identified and a method is proposed to automatically remove them from the outcome. By preventing any interaction between the physical and non-physical components, the proposed methodology drastically improves the robustness and reliability of the frame modal expansion modeling for thermoacoustics. The main numerical methodology relies on perturbation theory of symmetric eigenproblem and numerical approximation of the range of linear operators.

For more details on this work we refer to [15].

7.4 Computing WSBM marginals with Tensor-Train decomposition

The Weighted Stochastic Block Model (WSBM) is a statistical model for unsupervised clustering of individuals based on a pairwise distance matrix. The probabilities of group membership are computed as unary marginals of the joint conditional distribution of the WSBM, whose exact evaluation with brute force is out of reach beyond a few individuals. We propose to build an exact Tensor-Train (TT) decomposition of the multivariate joint distribution, from the SVD of each binary factor of a WSBM, which leads to variables separation. We present how to exploit this decomposition to compute unary and binary marginals. They are expressed without approximation as products of matrices involved in the TT decomposition. However, the implementation of the procedure faces several numerical challenges. First, the dimensions of the matrices involved grow faster than exponentially with the number of variables. We bypass this difficulty by using the format of TT-matrices. Second, the TT-rank of the products grows exponentially. Then, we use a numerical approximation of matrices product that guarantees a low TT-rank, the rounding. We compare the TT approach with two classical inference methods, the Mean-Field approximation and the Gibbs Sampler, on the problem of binary marginal inference for WSBM with various distances structures and up to fifty variables. The results lead to recommend the TT approach for its accuracy and reasonable computing time. Further researches should be devoted to the numerical difficulties for controlling the rank in rounding, to be able to deal with larger problems.

For more details on this work we refer to [22].

7.5 A filtered multilevel Monte Carlo method for estimating the expectation of cell-centered discretized random fields

In this work, we investigate the use of multilevel Monte Carlo (MLMC) methods for estimating the expectation of discretized random fields. Specifically, we consider a setting in which the input and output vectors of numerical simulators have inconsistent dimensions across the multilevel hierarchy. This requires the introduction of grid transfer operators borrowed from multigrid methods. By adapting mathematical tools from multigrid methods, we perform a theoretical spectral analysis of the MLMC

estimator of the expectation of discretized random fields, in the specific case of linear, symmetric and circulant simulators. We then propose filtered MLMC (F-MLMC) estimators based on a filtering mechanism similar to the smoothing process of multigrid methods, and we show that the filtering operators improve the estimation of both the small- and large-scale components of the variance, resulting in a reduction of the total variance of the estimator. Next, the conclusions of the spectral analysis are experimentally verified with a one-dimensional illustration. Finally, the proposed F-MLMC estimator is applied to the problem of estimating the discretized variance field of a diffusion-based covariance operator, which amounts to estimating the expectation of a discretized random field. The numerical experiments support the conclusions of the theoretical analysis even with non-linear simulators, and demonstrate the improvements brought by the F-MLMC estimator compared to both a crude MC and an unfiltered MLMC estimator.

For more details on this work we refer to [47].

7.6 Multilevel Surrogate-based Control Variates

Monte Carlo (MC) sampling is a popular method for estimating the statistics (e.g. expectation and variance) of a random variable. Its slow convergence has led to the emergence of advanced techniques to reduce the variance of the MC estimator for the outputs of computationally expensive solvers. The control variates (CV) method corrects the MC estimator with a term derived from auxiliary random variables that are highly correlated with the original random variable. These auxiliary variables may come from surrogate models. Such a surrogate-based CV strategy is extended here to the multilevel Monte Carlo (MLMC) framework, which relies on a sequence of levels corresponding to numerical simulators with increasing accuracy and computational cost. MLMC combines output samples obtained across levels, into a telescopic sum of differences between MC estimators for successive fidelities. In this paper, we introduce three multilevel variance reduction strategies that rely on surrogate-based CV and MLMC. MLCV is presented as an extension of CV where the correction terms devised from surrogate models for simulators of different levels add up. MLMC-CV improves the MLMC estimator by using a CV based on a surrogate of the correction term at each level. Further variance reduction is achieved by using the surrogate-based CVs of all the levels in the MLMC-MLCV strategy. Alternative solutions that reduce the subset of surrogates used for the multilevel estimation are also introduced. The proposed methods are tested on a test case from the literature consisting of a spectral discretization of an uncertain 1D heat equation, where the statistic of interest is the expected value of the integrated temperature along the domain at a given time. The results are assessed in terms of the accuracy and computational cost of the multilevel estimators, depending on whether the construction of the surrogates, and the associated computational cost, precede the evaluation of the estimator. It was shown that when the lower fidelity outputs are strongly correlated with the high-fidelity outputs, a significant variance reduction is obtained when using surrogate models for the coarser levels only. It was also shown that taking advantage of pre-existing surrogate models proves to be an even more efficient strategy.

For more details on this work we refer to [26]

7.7 Quadrature Rules in General Continuous Bayesian Networks : Discrete Inference without Discretization

Probabilistic inference in high-dimensional continuous or hybrid domains poses significant challenges, commonly addressed through discretization, sampling, or reliance on parametric assumptions. The drawbacks of these methods are well-known: inaccuracy, slow computational speeds or overly constrained models. This paper introduces a novel general inference algorithm designed for Bayesian networks featuring both discrete and continuous variables. The algorithm avoids the discretization of continuous densities into histograms by employing quadrature rules to compute continuous integrals and avoids the use of a parametric model by using orthogonal polynomials to represent the posterior density. Additionally, it preserves the computational efficiency of classical sum-product algorithms by using an auxiliary discrete Bayesian networks appropriately constructed to make continuous inference. Numerous experiments are conducted using either the conditional linear Gaussian model as a benchmark, or non-Gaussian models for greater generality. Our algorithm demonstrates significant improvements both in speed and accuracy when compared with existing methods

For more details on this work we refer to [29]

7.8 Material Parameter Estimation for a Viscoelastic Stenosis Model Using a Variational Autoencoder Inverse Mapper

Coronary artery disease, a prevalent condition often leading to heart attacks, may cause abnormal wall shear stresses near stenosed regions generating high frequent acoustic shear waves. In a previous study, a viscoelastic agarose gel was used to model the human tissue and it was shown that two material parameters of the gel could be estimated with a high certainty using a classical inverse problem. Given the high computational cost of traditional methods, this paper explores machine learning (ML) alternatives, particularly a Variational Autoencoder Inverse Mapper (VAIM). VAIM, previously successful in nuclear physics, uses neural networks to approximate forward and backward mappings and learn posterior parameter distributions. This paper validates previous research by generating data around ground truth values, demonstrating VAIM's ability to estimate two material parameters effectively. Further, it addresses realistic applications by training and testing on noisy data and generalizing findings across different intervals of signal damping.

For more details on this work we refer to [28]

7.9 Abstracting hierarchical methods for accelerating matrix-vector products on heterogeneous machine clusters: A unified framework for Fast Multipole Methods, \mathcal{H} -Matrices, and beyond

The Fast Multipole Method (FMM) is a renowned algorithm for accelerating the computation of interactions in N-body simulations and is recognized as one of the most influential algorithms of the 21st century. While commonly depicted as a tree-based approach, the FMM mathematically represents a block matrix-vector product, where hierarchical extra-diagonal block operations are compressed analytically, echoing the literature on \mathcal{H} -matrices [46] where those are compressed algebraically. However, few attempts have been made to bridge the gap between FMM and \mathcal{H} matrices ([61], [63]). The goal of the PhD aims to abstract and unify hierarchical methods, including the FMM, \mathcal{H} -matrices, and their flat counterparts (Block Low-Rank matrices [71] for instance), especially for computing fast matrix-vector products.

In a paper currently being drafted, we propose an abstraction of hierarchical methods that consists of four main components: (1) a partitioning scheme to decompose degrees of freedom and organize the matrix into a block structure, (2) an admissibility criterion to predict which blocks to compress and which to leave intact, (3) a method for compressing admissible blocks (analytical, algebraic, or arithmetic...), and (4) a strategy to manage bases. The objective is to develop a single algorithm that could accommodate all hierarchical methods and adapt to various physical problem sizes and target computer characteristics.

This work is part of Antoine Gicquel's PhD thesis and is carried out in the context of the Diwina ANR project.

7.10 Using the Fast Multipole Method to Accelerate Matrix-Vector Products

In electrical engineering, solving low-frequency Maxwell's equations is crucial for modeling devices with ferromagnetic materials and large air volumes, typically due to voids and gaps. The Magnetostatic Moment Method (MMM), based on volume integral equations, offers an efficient alternative to the Finite Element Method (FEM) in such cases. The MMM's advantage lies in not requiring the meshing of air regions, significantly reducing the number of unknowns. However, the MMM results in linear systems with dense matrices, leading to computational challenges in terms of memory and execution time for large problems. To address this issue, we consider a low-rank approximation of the matrix using the Fast Multipole Method (FMM). This hierarchical algorithm, which uses a tree-based domain partitioning, can reduce the complexity of matrix-vector product calculations from $O(N^2)$ to $O(N)$. Our work focuses on adapting a black-box FMM variant, available in the C++ ScalFMM library, to accelerate matrix-vector products within the MMM framework. Additionally, we developed a parallel, shared memory-based version to further enhance computational efficiency. Our numerical results demonstrate the effectiveness

of this enhanced FMM-based matrix-vector product in a simplified application modeling a low-frequency antenna.

For more details on this work we refer to [33]

7.11 Optimizing multigrid efficiency using machine learning techniques

Our research explores the integration of machine learning into multigrid solvers, focusing on the similarities between UNet architectures and multigrid cycles. While traditional geometric multigrid methods rely on fixed stencils for smoothing and prolongation/restriction steps, these operations can be reinterpreted as trainable parameters within a neural network framework. During the project, we developed a flexible multigrid cycle implementation in PyTorch, where we replace matrix-vector multiplications with stencils by 2D convolutions. This approach enables us to train stencil values using gradient-based optimization techniques. The multigrid solver is thus effectively transformed into a hybrid learning-based solver. Our work investigates various optimization strategies to improve performance and generalization of the multigrid cycle, where we train one or more components at the same time. We aim to identify which parameters and methods have the greatest impact on efficiency, both in terms of computational cost and solution quality.

7.12 Comparison of multigrid and machine learning-based Poisson solvers

We present a comprehensive comparison of the multigrid method and the UNet architecture for solving Poisson's equation. Those two methods show a lot of similarity and also have the very interesting characteristic that their solving time should scale linearly with the number of mesh nodes. Nevertheless, for Poisson's equation, an analysis of the number of floating-point operations demonstrates that the multi-grid V-Cycle should be faster than the UNet. We have realized a practical comparison of the two methods solving time for different number of mesh nodes on the same computation nodes using GPU.

For more details on this work we refer to [27]

7.13 Statistical machine learning techniques to predict the rank in \mathcal{H} -matrix computation

The efficient solution of large, dense linear systems arising from the Boundary Element Method (BEM) for integral equations on complex geometries is nowadays often addressed using Hierarchical matrix (\mathcal{H} -matrix) techniques. The main challenges are twofold: first, determining the hierarchical block low-rank structure of the matrix—i.e., identifying which blocks have low rank and can be advantageously stored as the product of rectangular matrices; second, estimating the actual ranks of these blocks. In this work, we explore supervised statistical machine learning techniques to tackle these challenges. More specifically, we train models based on decision trees (specifically, random forests or gradient-boosted trees) on synthetic datasets generated from simple geometries and validate their generalization capabilities on 3D aircraft meshes.

This work is part of Théo Briquet's PhD.

7.14 Convergence analysis of overlapping domain decomposition preconditioners for nonlinear problems

Numerical simulations of nonlinear partial differential equations often involve solving large nonlinear systems, for which Newton's method is widely employed due to its rapid convergence near the solution. However, its performance can deteriorate in the presence of strong nonlinearities or poor initial guesses. Nonlinear overlapping domain decomposition methods, such as RASPEN [55] and Substructured RASPEN (SRASPEN) [51], have shown to address these challenges effectively. While SRASPEN reduces the problem size by restricting computations to a substructure, it suffers from a loss of information outside the substructure, leading to additional inner subdomain iterations.

In this work, we analyze the convergence of RASPEN, demonstrating how domain decomposition improves the convergence rate of Newton's method by highlighting the impact of substructuring on the global domain error contraction. Furthermore, we propose an inexpensive adjustment to SRASPEN

to mitigate the loss of global information, thereby reducing computational cost and improving overall efficiency. Numerical experiments show the computational performance of the improved SRASPEN, establishing it as a robust approach for solving large-scale nonlinear systems.

This work is part of Ettaouchi El Mehdi's PhD thesis and is carried out in collaboration with Nicolas Tardieu (EDF).

7.15 Low-rank tensor solver for magnetostatic problems for electric power applications

The development of electrical and electromechanical devices requires solving Maxwell's equations, which is challenging due to their complexity. Volumetric integral methods (VIM) are advantageous for low-frequency problems as they eliminate the need to mesh air regions. However, they result in dense matrices, leading to high computational costs. Compression techniques, such as fast multipole methods (FMM) and \mathcal{H} -Matrix, have mitigated these issues by reducing complexity.

This thesis introduces low-rank tensors, specifically Tensor-Train, to further compress systems and accelerate computations by transforming problems into a tensor framework without constructing full tensors. We focused on two VIM methods: the Magnetic Moment Method and the Scalar Potential Method, both of which were adapted to a tensor format. Additionally, we conducted comparisons of these tensor-based methods with \mathcal{H} -Matrix and FMM approaches on simple test cases to evaluate their performance and efficiency.

To solve the compressed linear systems, two approaches were explored: solvers based on optimization techniques and classical iterative solvers, such as Krylov subspace methods. The Tensor-Train approach demonstrated better performance in terms of both compression and solver computation time, making it particularly efficient for large-scale problems. While the methods have been tested on simple problems to validate their effectiveness, future work will focus on applying them to more complex geometries to further evaluate their scalability and robustness.

This work is part of Amine Zekri's PhD thesis and is carried out in the context of the TensoVim ANR project.

7.16 Intra-node performance

As part of the EoCoE III project, this work aims to enable robust and high-performance solutions for large sparse linear systems on modern platforms characterized by increasing complexity and resource heterogeneity. The focus is on optimizing the parallelization of domain decomposition algorithms within the Composyx software framework, employing both direct and iterative methods to enhance scalability and accuracy. These algorithms rely on several state-of-the-art libraries, including BLAS/LAPACK for dense linear algebra operations, direct sparse solvers for exact subproblem resolution, and partitioners for efficient domain decomposition. Achieving seamless integration and optimal utilization of these libraries is critical for ensuring computational efficiency. Our current efforts target homogeneous multi-core architectures, aiming to optimize all algorithmic stages for maximal performance. For heterogeneous platforms, particularly CPU+GPU nodes, we propose a task-based approach that decomposes the computational workflow into fine-grained tasks. This strategy facilitates load balancing and efficient resource allocation, leveraging the complementary strengths of CPUs and GPUs. In this context, we are developing innovative techniques to express and exploit parallelism, enabling the algorithms to scale effectively on exascale systems. By addressing 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.

This work is conducted in collaboration with IRIT.

8 Partnerships and cooperations

Participants: All team members.

8.1 European initiatives

8.1.1 H2020 projects

EoCoE-3

Title: Energy oriented Centre of Excellence for computer applications

Duration: 2024-2026

Coordinator: CEA

Inria coordinator: Bruno Raffin

CONCACE contact: Emmanuel Agullo

Partners:

- AGENZIA NAZIONALE PER LE NUOVE TECNOLOGIE, L'ENERGIA E LO SVILUPPO ECONOMICO SOSTENIBILE (Italy)
- BARCELONA SUPERCOMPUTING CENTER - CENTRO NACIONAL DE SUPERCOMPUTACION (Spain)
- CENTRE EUROPEEN DE RECHERCHE ET DE FORMATION AVANCEE EN CALCUL SCIENTIFIQUE (France)
- CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE CNRS (France)
- COMMISSARIAT A L ENERGIE ATOMIQUE ET AUX ENERGIES ALTERNATIVES (France)
- CONSIGLIO NAZIONALE DELLE RICERCHE (Italy)
- FORSCHUNGSZENTRUM JULICH GMBH (Germany)
- FRAUNHOFER GESELLSCHAFT ZUR FOERDERUNG DER ANGEWANDTEN FORSCHUNG E.V. (Germany)
- MAX-PLANCK-GESELLSCHAFT ZUR FORDERUNG DER WISSENSCHAFTEN EV (Germany)
- RHEINISCH-WESTFAELISCHE TECHNISCHE HOCHSCHULE AACHEN (Germany)
- UNIVERSITA DEGLI STUDI DI ROMA TORVERGATA (Italy)
- UNIVERSITA DEGLI STUDI DI TRENTO (Italy)
- UNIVERSITE LIBRE DE BRUXELLES (Belgium)
- UNIVERSITE PARIS-SUD (France)

Inria contact: Bruno Raffin (Datamove)

Summary: The Concace team (Inria, Cerfacs) participates in the Energy-oriented Centre of Excellence (EoCoE-III), starting in January 2024. The project applies cutting-edge exascale computational methods in its mission to accelerate the transition to the production, storage and management of clean, decarbonized energy. EoCoE-III is anchored in the High Performance Computing (HPC) community and targets research institutes and key commercial players who develop and enable energy-relevant numerical models to be run on exascale supercomputers, demonstrating their benefits for the net-zero energy transition. The project will draw on the experience of two successful previous projects EoCoE-I and -II, where a large set of diverse computer applications from four such energy domains achieved significant efficiency gains thanks to a multidisciplinary expertise in applied mathematics and supercomputing. EoCoE-III channels its efforts into 5 exascale lighthouse applications in the low-carbon sectors of Energy Materials, Water, Wind and Fusion. This multidisciplinary effort will harness innovations in computer science and mathematical algorithms within a tightly integrated co-design approach to overcome performance bottlenecks and to anticipate HPC hardware developments. A world-class consortium of 16 complementary partners forms a unique network of expertise in energy science, scientific computing and HPC, including 3 leading European supercomputing centres.

8.1.2 Other european programs/initiatives

Title: High Performance Spacecraft Plasma Interaction Software

Duration: 2022 - 2024

Funding: ESA

Coordinator: Sébastien Hess (ONERA)

Concace contact: Olivier Couland and Luc Giraud

Partners: • Airbus DS

- Artenum
- ONERA

Summary: Controlling the plasma environment of satellites is a key issue for the nation in terms of satellite design and propulsion. Three-dimensional numerical modelling is thus a key element, particularly in the preparation of future space missions. The SPIS code is today the reference in Europe for the simulation of these phenomena. The methods used to describe the physics of these plasmas are based on the representation of the plasma by a system of particles moving in a mesh (here unstructured) under the effect of the electric field which satisfies the Poisson equation. ESA has recently shown an interest in applications requiring complex 3D calculations, which may involve several tens of millions of cells and several tens of billions of particles, and therefore in a highly parallel and scalable version of the SPIS code.

8.2 National initiatives

MAMBO

Duration: 2018 – 2022

CONCACE contact: Guillaume Sylvand

Funding: DGAC

Partners:

- CEA
- Inria
- CNRS

Summary: MAMBO ("Méthodes Avancées pour la Modélisation du Bruit moteur et aviOn") is a project funded by the DGAC, bringing together 21 academic and industrial partners from France and Europe, including Airbus, Cerfacs, and Inria. For Inria, the key challenge of this project lies in addressing new research problems and developing innovative numerical methods by enhancing the capabilities of its software, particularly in the field of high-performance computing.

PEPR Numpex

Duration: 2018 – 2022

CONCACE contact: Emmanuel Agullo, Luc Giraud

Funding: ANR

Partners:

- CEA
- Inria
- CNRS

Summary: NumPEX is a French program dedicated to Exascale: High-performance computing (HPC), high-performance data analytics (HPDA), and Artificial Intelligence (AI) pose significant challenges across scientific, societal, economic, and ethical realms. These technologies, including modeling and data analysis, are crucial decision support tools addressing societal issues and competitiveness in French research and development. Digital resources, essential across science and industry, demand high-performance hardware. HPC enables advanced modeling, while HPDA handles heterogeneous and massive data. The solution to exploding demand is the upcoming “exascale” computers, a new generation with extraordinary capabilities.

In this context, the French Exascale program NumPEX aims at designing and developing the software components that will equip future exascale machines. NumPEX will deliver Exascale-grade numerical methods, softwares, and training, allowing France to remain one of the leaders in the field. It will contribute to take bridging the gap between cutting-edge software development and application domains to prepare the major scientific and industrial application codes to fully exploit the capabilities of these machines. Application domains of the NumPEX program include, but are not limited to, weather forecasting and climate, aeronautics, automotive, astrophysics, high energy physics, material science, energy production and management, biology and health.

NumPEX is organized in 7 scientific pillar projects, we are directly involved in two of them namely:

Exa-MA : Methods and Algorithms for Exascale;

Exa-Soft : HPC softwares and tools.

TensorVIM

Duration: 2023 – 2026

Coordinator: LAPLACE

CONCACE contact: Olivier Coulaud

Funding: ANR

Partners:

- Inria
- LAPLACE
- G2ELaB

Summary: The aim of this project is to develop high-performance computational tools for the rapid implementation of low-frequency electromagnetic simulations for electrical applications. We consider an approach based on volume integral methods using low-rank approximations. Instead of using classical compression techniques such as the fast multipole method or the hierarchical matrix approach, we propose to investigate the use of low-rank tensors to accelerate the computation of the solution of the linear system. The tools developed will be used for the modeling of various devices (PCB modeling, Electrical Machines) with the main goal of improving their energy performance.

Maturation

Title: MAssively parallel sparse grid PIC algorithms for low TemperatURE plAsmas SimulaTIONS

Duration: 2023 – 2026

Coordinator: Laurent Garrigues (Laplace)

CONCACE contact: Luc Giraud

Funding: ANR

Partners:

- Laplace Lab
- IMT

Summary: The simulation under real conditions of partially magnetized low temperature plasmas by Lagrangian approaches, though using powerful Particle-In-Cell (PIC) techniques supplemented with efficient high-performance computing methods, requires considerable computing resources for large plasma densities. This is explained by two main limitations. First, stability conditions that constrain the numerical parameters to resolve the small space and time scales. These numerical parameters are the mesh size of the grid used to compute the electric field and the time step between two consecutive computations. Second, PIC methods rely on a sampling of the distribution function by numerical particles whose motion is time integrated in the self-consistent electric field. The PIC algorithm remains close to physics and offers an incomparable efficiency with regard to Eulerian methods, discretizing the distribution function onto a mesh. It is widely and successfully operated for the discretization of kinetic plasma models for more than 40 years. Nonetheless, to spare the computational resources, the number of numerical particles is limited compared to that of the physical particles. Inherent to this “coarse” sampling, PIC algorithms produce numerical approximations prone to statistical fluctuations that vanish slowly with the mean number of particles per cell. The mesh accessible on typical high performance computing machines may 10^9 cells, which brings the mesh size close to the scale of the physics, but the mean number of numerical particles in each cell shall be limited, to mitigate the memory footprint as well as the computational time. A breakthrough is therefore necessary to reduce the computational resources by orders of magnitude and make possible the use of explicit PIC method for large scale and/or densities for 3D computations.

This is the issue addressed within the MATURATION project aiming at introducing a new class of PIC algorithms with an unprecedented computational efficiency, by analyzing and improving, parallelizing and optimizing as well as benchmarking, in the demanding context of partially magnetized low temperature plasmas through 2D large scale and 3D computations, a method recently proposed in the literature, based on a combination of sparse grid techniques and PIC algorithm.

Diwina

Title: Magnetic Digital Twins for Spintronics : nanoscale simulation platform

Duration: 2023 – 2026

Coordinator: Institut Neel

CONCACE contact: Olivier Coulaud

Funding: ANR

Partners:

- CMAP, Institut Neel, Inria, SPINTEC

Summary: The DiWiNa project aims at developing a unified open-access platform for spintronic numerical twins, ie, codes for micromagnetic/spintronic simulations with sufficiently-high reliability and speed so that they can be trusted and used as reality. The simulations will be bridged to the advanced microscopy techniques used by the community, through plugins to convert the statics or time-resolved 3D vector- fields into contrast maps for the various techniques, including their

experimental transfer functions. To achieve this, we bring together experts from different disciplines to address the various challenges: spintronics for the core simulations, mathematics for trust, algorithmics for speed, experimentalists for the bridge with microscopy. Practical work consists of checking the time-integration stability of spintronic torque involved in the dynamics when implemented in the versatile finite-element framework, improve the calculation speed through advanced libraries, build the bridge with microscopies through rendering tools, and encapsulate these three key ingredients into a user-friendly Python ecosystem. Through open-access and versatile user-friendly encapsulation, we expect that this platform is suited to serve the needs of the entire physics and engineering community of spintronics. The platform will be unique in its features, ranging from simulation to the direct and practical comparison with experiments. It will contribute to reduce considerably the number of experimental screening for the faster development of new spintronic devices, which are expected to play a key role in energy saving.

9 Dissemination

Participants: All permanent team members.

9.1 Promoting scientific activities

9.1.1 Scientific events: organisation

Member of the organizing committees

- Carola Kruse and Paul Mycek are members of the organising committee of the “Sparse Days 2024”

9.1.2 Scientific events: selection

Member of the conference program committees

- PDSEC: Olivier Coulaud, Luc Giraud,
- Luc Giraud is member of the Gene Golub SIAM Summer School. The twelfth Gene Golub SIAM Summer School was entitled “Iterative and Randomized Methods for Large-Scale Inverse Problems”, Campus of the Escuela Politécnica Nacional, Ecuador, 22 July to 2 August 2024

Co-chair of conference proceedings

- ISC-HPC 2024: Carola Kruse

Reviewer

- ISC-HPC 2024: Carola Kruse for Birds of a feather submissions

9.1.3 Journal

Reviewer - reviewing activities BIT, IMA Journal of Numerical Analysis, Nature scientific report, SIAM/ASA journal on uncertainty quantification ...

9.1.4 Invited talks

- A journey on subspace methods for the solution of sequences of linear systems by Luc Giraud, Emmanuel Agullo, Olivier Coulaud, Martina Iannacito, Gilles Marait, Yanfei Xiang; at International Conference on Mathematics and Decision, December 17–20, 2024; Mohammed VI Polytechnic University, Morocco.

9.1.5 Scientific expertise

- Luc Giraud is
 - member of the board on Modelization, Simulation and data analysis of the Competitiveness Cluster for Aeronautics, Space and Embedded Systems.
 - member of the scientific council of the ONERA Lab LMA2S (Laboratoire de Mathématiques Appliquées à l'Aéronautique et au Spatial).
 - member of the scientific council of GDR Calcul.
 - referee for Czech Science Foundation proposal.
- Carola Kruse referee for Icelandic Research Fund proposals.
- Guillaume Sylvand is
 - expert in Numerical Simulation and HPC at Airbus.
 - member of the scientific council of the ORAP.

9.1.6 Research administration

- Emmanuel Agullo is member of the CDT (Technological Development Commission) at inria Centre at the Bordeaux University.
- Luc Giraud is techniques pilot for the expert group for the evaluation of French research entities (UMRs and EAs) relatively to the protection of scientific and technological properties (PPST) on information and communication sciences and technologies (STIC).

9.2 Teaching - Supervision - Juries

9.2.1 Teaching

- Post graduate level/Master:
 - E. Agullo: Operating systems 24h at Bordeaux University ; Dense linear algebra kernels 8h, Numerical algorithms 30h at Bordeaux INP (ENSEIRB-MatMeca).
 - O. Coulaud: Paradigms for parallel computing 8h, Introduction to Tensor methods 6 h at Bordeaux INP (ENSEIRB-MatMeca).
 - L. Giraud: Introduction to intensive computing and related programming tools 20h, INSA Toulouse; Advanced numerical linear algebra 10h, ENSEEIHT Toulouse.
 - C. Kruse: Adavanced topics in numerical linear algebra, 10h, FAU Erlangen; Méthodes Itératives en Algèbre Linéaire, 14h, ENSEEIHT Toulouse.
 - P. Mycek: Multifidelity methods 14h, ModIA (cursus en alternance, INSA/N7), Toulouse.
 - L. Giraud, C. Kruse, P. Mycek: supervision of INSA applied math students (4A) on “initiation to reasearch” projects (2x12h)

9.2.2 Supervision

- PhD in progress: Alexandre Malhene; Abstraction of subspace methods in numerical linear algebra; started October 2024, E. Agullo, L. Giraud.
- PhD in progress: Hugo Dodelin; Abstraction of parallel execution models; started October 2024, E. Agullo, O. Coulaud.
- PhD in progress: Théo Briquet; machine learning techniques for rank prediction of \mathcal{H} -matrices; started October 2023, L. Giraud, P. Mycek, G. Sylvand.

- PhD in progress: Mehdi El Ettaouchi; nonlinear domain decomposition techniques in geosciences; started March 2023, L. Giraud, C. Kruse, N. Tardieu (EDF).
- PhD in progress: Antoine Gicquel; Acceleration of the matrix-vector product by the fast multipole method for heterogeneous machine clusters; started Nov. 2023, O. Coulaud, B. Bramas.
- PhD in progress: Andrea Lagardère; Méthode Quasi-Trefftz Couplée pour l'Aéroacoustique; started April 2024, G Sylvand, S. Tordeux.
- PhD in progress: Amine Zekri ; Low-rank tensor solver for magnetostatic problems for electric power applications, started October 2023; O. Coulaud, J.R. Poirier

9.2.3 Juries

PhD defense

- Jérémy Briant, "Méthodes de Monte Carlo multi-niveaux pour l'estimation de paramètres statistiques de champs discrétisés en géosciences", referees: Olivier Le Maître, Laurent Debreu; members: Serge Gratton, Anthony Weaver, Clémentine Prieur, Mohamed Reda El Amri, Selime Gürol; invited: Paul Mycek, Ehouarn Simon; École doctorale Mathématiques, informatique et télécommunications (Toulouse).
- Marc Chung To Sang, "Electronic transport and secondary emission in a Hall thruster"; referees: Aaron Knoll, Kentaro Hara; members: Fabrice Deluzet, Gwenael Fubiani, Laurent Garriques, Luc Giraud, Sédina Tsikata; Université de Toulouse; spécialité: Génie Electrique, Electronique, Télécommunications et Santé : du système au nanosystème.
- Matthieu Robeyns, "Mixed precision algorithms for low-rank matrix and tensor approximations"; referees: Anthony Nouy, Rio Yokota; members: Emmanuel Agullo, Sylvie Boldo, Mariya Ishteva, Marc Baboulin, Oguz Kaya, Théo Mary; Université Paris-Saclay; spécialité: Informatique et sciences du numérique.
- Johan Valentin, "Couplage multi-fidélité de particules tourbillonnaires avec lignes portantes et méthode Eulérienne pour la simulation d'écoulements aéronautiques 3D visqueux", referees: Annie Leroy, Iraj Mortazavi; members: Elie Rivoalen, Grégory Pinon, Luis Bernardos, Paul Mycek, Ivan Delbende, Iraj Mortazavi, Annie Leroy, Chloé Mimeau; INSA Rouen Normandie; École Doctorale Physique, Sciences de l'Ingénieur, Matériaux, Énergie.

10 Scientific production

10.1 Major publications

- [1] E. Agullo, B. Bramas, O. Coulaud, E. Darve, M. Messner and T. Takahashi. 'Task-Based FMM for Multicore Architectures'. In: *SIAM Journal on Scientific Computing* 36.1 (2014), pp. 66–93. DOI: [10.1137/130915662](https://doi.org/10.1137/130915662). URL: <https://hal.inria.fr/hal-00911856>.
- [2] E. Agullo, A. Buttari, A. Guermouche, J. Herrmann and A. Jego. 'Task-based parallel programming for scalable matrix product algorithms'. In: *ACM Transactions on Mathematical Software* (2023). URL: <https://hal.science/hal-03936659>.
- [3] E. Agullo, L. Giraud and L. Poirel. 'Robust preconditioners via generalized eigenproblems for hybrid sparse linear solvers'. In: *SIAM Journal on Matrix Analysis and Applications* 40.2 (2019), pp. 417–439. DOI: [10.1137/17M1153765](https://doi.org/10.1137/17M1153765). URL: <https://hal.inria.fr/hal-02074474>.
- [4] E. Bourne, P. Leleux, K. Kormann, C. Kruse, V. Grandgirard, Y. Güçlü, M. J. Kühn, U. Rüde, E. Sonnendrücker and E. Zoni. *Solver comparison for Poisson-like equations on tokamak geometries*. 23rd Sept. 2022. URL: <https://cea.hal.science/cea-03786723>.
- [5] B. Bramas, O. Coulaud and G. Sylvand. 'Time-domain BEM for the wave equation on distributed-heterogeneous architectures: A blocking approach'. In: *Parallel Computing* 49 (July 2015), pp. 66–82. DOI: [10.1016/j.parco.2015.07.005](https://doi.org/10.1016/j.parco.2015.07.005). URL: <https://hal.inria.fr/hal-01200727>.

- [6] F. Casenave, A. Ern and G. Sylvand. ‘Coupled BEM-FEM for the convected Helmholtz equation with non-uniform flow in a bounded domain’. In: *Journal of Computational Physics* 257.A (15th Jan. 2014), pp. 627–644. DOI: [10.1016/j.jcp.2013.10.016](https://doi.org/10.1016/j.jcp.2013.10.016). URL: <https://hal.science/hal-00805477>.
- [7] S. Cools, E. F. Yetkin, E. Agullo, L. Giraud and W. Vanroose. ‘Analyzing the Effect of Local Rounding Error Propagation on the Maximal Attainable Accuracy of the Pipelined Conjugate Gradient Method’. In: *SIAM Journal on Matrix Analysis and Applications* 39.1 (13th Mar. 2018), pp. 426–450. DOI: [10.1137/17M1117872](https://doi.org/10.1137/17M1117872). URL: <https://hal.inria.fr/hal-01753411>.
- [8] V. Darrigrand, A. Dumitras, C. Kruse and U. Rude. ‘Inexact inner–outer Golub–Kahan bidiagonalization method: A relaxation strategy’. In: *Numerical Linear Algebra with Applications* (12th Dec. 2022). DOI: [10.1002/nla.2484](https://doi.org/10.1002/nla.2484). URL: <https://hal.inria.fr/hal-03960074>.
- [9] D. A. Di Pietro, P. Matalon, P. Mycek and U. Rude. ‘High-order multigrid strategies for HHO discretizations of elliptic equations’. In: *Numerical Linear Algebra with Applications* (22nd June 2022). DOI: [10.1002/nla.2456](https://doi.org/10.1002/nla.2456). URL: <https://hal.science/hal-03531293>.
- [10] A. Franc, P. Blanchard and O. Coulaud. ‘Nonlinear mapping and distance geometry’. In: *Optimization Letters* 14.2 (2020), pp. 453–467. DOI: [10.1007/s11590-019-01431-y](https://doi.org/10.1007/s11590-019-01431-y). URL: <https://hal.inria.fr/hal-02124882>.
- [11] L. Giraud, Y.-F. Jing and Y. Xiang. ‘A block minimum residual norm subspace solver with partial convergence management for sequences of linear systems’. In: *SIAM Journal on Matrix Analysis and Applications* 43.2 (2022), pp. 710–739. DOI: [10.1137/21m1401127](https://doi.org/10.1137/21m1401127). URL: <https://hal.inria.fr/hal-03546496>.
- [12] Y. Jang, J. Bonelle, C. Kruse, F. Hulsemann and U. Ruede. *Fast Linear Solvers for Incompressible CFD Simulations with Compatible Discrete Operator Schemes*. 27th Apr. 2023. URL: <https://hal.science/hal-04087358>.
- [13] M. Odunlami, V. Le Bris, D. Begue, I. Baraille and O. Coulaud. ‘A-VCI: A flexible method to efficiently compute vibrational spectra’. In: *Journal of Chemical Physics* 146.21 (7th June 2017). DOI: [10.1063/1.4984266](https://doi.org/10.1063/1.4984266). URL: <https://hal.inria.fr/hal-01534134>.

10.2 Publications of the year

International journals

- [14] V. Acary, P. Armand, H. Minh Nguyen and M. Shpakovych. ‘Second order cone programming for frictional contact mechanics using interior point algorithm’. In: *Optimization Methods and Software*. Special issue honoring Oleg Burdakov 39.3 (2024), pp. 634–663. DOI: [10.1080/10556788.2023.2296438](https://doi.org/10.1080/10556788.2023.2296438). URL: <https://hal.science/hal-03913568>.
- [15] M. Cances, L. Giraud, M. Bauerheim, L. Gicquel and F. Nicoud. ‘Robustness and reliability of state-space, frame-based modeling for thermoacoustics’. In: *Journal of Computational Physics* 520 (Jan. 2025), p. 113472. DOI: [10.1016/j.jcp.2024.113472](https://doi.org/10.1016/j.jcp.2024.113472). URL: <https://hal.science/hal-04734953> (cit. on p. 12).
- [16] F. Cappello, S. Di, R. Underwood, D. Tao, J. Calhoun, Y. Kazutomo, K. Sato, A. Singh, L. Giraud, E. Agullo, X. Yepes, M. Acosta, S. Jin, J. Tian, F. Vivien, B. Zhang, K. Sano, T. Ueno, T. Grutzmacher and H. Anzt. ‘Multifacets of lossy compression for scientific data in the Joint-Laboratory of Extreme Scale Computing’. In: *Future Generation Computer Systems* (13th June 2024). DOI: [10.1016/j.future.2024.05.022](https://doi.org/10.1016/j.future.2024.05.022). URL: <https://inria.hal.science/hal-04618060> (cit. on p. 11).
- [17] B. Gobe, J. Saucourt, M. Shpakovych, D. Helbert, A. Desfarges-Berthelemot and V. Kermene. ‘Retrieving the complex transmission matrix of a multimode fiber by machine learning for 3D beam shaping’. In: *Journal of Lightwave Technology* (2024), pp. 1–8. DOI: [10.1109/JLT.2024.3373689](https://doi.org/10.1109/JLT.2024.3373689). URL: <https://hal.science/hal-04516609>.

International peer-reviewed conferences

- [18] L. Giraud, E. Agullo, O. Coulaud, M. Iannacito, G. Marait and Y. Xiang. ‘A journey on subspace methods for the solution of sequences of linear systems’. In: *International conference on Mathematics and decision*. Rabat, Morocco, Dec. 2024. URL: <https://inria.hal.science/hal-04854993>.
- [19] B. Gobé, J. Saucourt, M. Shpakovych, G. Maulion, D. Helbert, D. Pagnoux, A. Desfarges-Berthelemt and V. Kermène. ‘Machine learning method to measure the transmission matrix of a multimode optical fiber without reference beam for 3D beam tailoring’. In: *Proceedings of SPIE. Photonic West 2024 : Laser Resonators, Microresonators, and Beam Control XXVI*. Vol. 12871. Laser Resonators, Microresonators, and Beam Control XXVI. San Francisco, CA, United States: SPIE, 12th Mar. 2024, 128710F. DOI: [10.1117/12.3001422](https://doi.org/10.1117/12.3001422). URL: <https://hal.science/hal-04245837>.

Scientific book chapters

- [20] R. Espoey, L. Brevault, M. Balesdent, S. Ricci, P. Mycek and G. Arnoult. ‘Overview and comparison of reliability analysis techniques based on multifidelity Gaussian processes’. In: *Developments in Reliability Engineering*. Elsevier, 2024, pp. 731–785. DOI: [10.1016/B978-0-443-13242-1.00015-1](https://doi.org/10.1016/B978-0-443-13242-1.00015-1). URL: <https://hal.science/hal-04878717>.

Doctoral dissertations and habilitation theses

- [21] P. Mycek. ‘Hierarchical methods for deterministic and stochastic partial differential equations’. Université de Bordeaux, 8th Feb. 2024. URL: <https://hal.science/tel-04521927> (cit. on p. 9).

Reports & preprints

- [22] M. A. Abouabdallah, O. Coulaud, N. Peyrard and A. Franc. *Computing WSBM marginals with Tensor-Train decomposition*. 15th Jan. 2024. URL: <https://hal.inrae.fr/hal-04394024> (cit. on p. 12).
- [23] C. Acary-Robert, E. Agullo, L. Courtès, M. Felšöci, K. Hinsén, A. Isaac, O. Lünsdorf, P. Prins, S. Tournier, P. Virouleau and R. Wurmus. *Guix-HPC Activity Report 2022–2023*. Inria Bordeaux - Sud Ouest, Feb. 2024, pp. 1–32. URL: <https://inria.hal.science/hal-04500140>.
- [24] O. Beaumont, R. Bouzel, L. Eyraud-Dubois, E. Korkmaz, L. Lima Pilla and A. van Kempen. *Approximation Algorithms for Scheduling with/without Deadline Constraints where Rejection Costs are Proportional to Processing Times*. 15th Oct. 2024. URL: <https://hal.science/hal-04745701>.
- [25] M. Destouches, P. Mycek, S. Gürol, A. Weaver, S. Gratton and E. Simon. *Multilevel Monte Carlo methods for ensemble variational data assimilation*. 3rd Dec. 2024. DOI: [10.5194/egusphere-2024-3628](https://doi.org/10.5194/egusphere-2024-3628). URL: <https://hal.science/hal-04869802>.
- [26] M. R. El Amri, P. Mycek, S. Ricci and M. De Lozzo. *Multilevel Surrogate-based Control Variates*. 2024. URL: <https://hal.science/hal-04129561> (cit. on p. 13).
- [27] H. Godé, C. KRUSE, R. Angersbach, H. Köstler, M. Bauerheim and U. Råde. *Comparison of multigrid and machine learning-based Poisson solvers*. Aug. 2024. URL: <https://hal.science/hal-04689025> (cit. on p. 15).
- [28] C. Kruse, M. Sosonkina, M. Fayaz Bin Hossen and Y. Li. *Material Parameter Estimation for a Viscoelastic Stenosis Model Using a Variational Autoencoder Inverse Mapper*. July 2024. URL: <https://hal.science/hal-04700426> (cit. on p. 14).
- [29] M. Lasserre, R. Lebrun and P.-H. Wuillemin. *Quadrature Rules in General Continuous Bayesian Networks : Discrete Inference without Discretization*. 10th Feb. 2024. URL: <https://hal.science/hal-04495263> (cit. on p. 14).
- [30] A. Tremodeux, E. Agullo, A. Benoit, L. Giraud, T. Herault and Y. Robert. *Fault-tolerant numerical iterative algorithms at scale*. RR-9567. Inria Lyon, Jan. 2025. URL: <https://inria.hal.science/hal-04872041>.

- [31] N. Venkovic, P. Mycek, O. L. Maître and L. Giraud. *Preconditioners based on Voronoi quantizers of random variable coefficients for stochastic elliptic partial differential equations*. 12th Mar. 2024. URL: <https://hal.science/hal-04790361> (cit. on p. 12).
- [32] Y. Xiang. *Unsupervised convolution neural operator preconditioning for the solution of some heterogeneous fluid PDEs*. Inria Centre at the University of Bordeaux, 14th Jan. 2025. URL: <https://inria.hal.science/hal-04886933>.

Other scientific publications

- [33] A. Gicquel and O. Coulaud. *Using the Fast Multipole Method to Accelerate Matrix-Vector Products: Application to the Magnetostatic Moment Method*. Paris, France, May 2024. URL: <https://inria.hal.science/hal-04588314> (cit. on p. 15).
- [34] M. Iannacito, O. Coulaud and L. Giraud. *Orthogonalization schemes in tensor train format*. Paris, France, 13th May 2024. URL: <https://inria.hal.science/hal-04587668>.
- [35] Y. Xiang, L. Giraud, P. Mycek, M. Shpakovych and C. Kruse. *Neural operator preconditioning for accelerating the solution of the parametric Helmholtz equations*. Paris, France, 13th May 2024. URL: <https://inria.hal.science/hal-04581208>.

10.3 Cited publications

- [36] E. Agullo, O. Aumage, B. Bramas, O. Coulaud and S. Pitoiset. ‘Bridging the gap between openMP and task-based runtime systems for the fast multipole method’. In: *IEEE Transactions on Parallel and Distributed Systems* 28.10 (2017). DOI: [10.1109/TPDS.2017.2697857](https://doi.org/10.1109/TPDS.2017.2697857) (cit. on p. 7).
- [37] E. Agullo, B. Bramas, O. Coulaud, E. Darve, M. Messner and T. Takahashi. ‘Task-based FMM for heterogeneous architectures’. In: *Concurrency and Computation: Practice and Experience* 28.9 (June 2016), pp. 2608–2629. DOI: [10.1002/cpe.3723](https://doi.org/10.1002/cpe.3723). URL: <http://doi.wiley.com/10.1002/cpe.3723> (cit. on p. 7).
- [38] E. Agullo, B. Bramas, O. Coulaud, E. Darve, M. Messner and T. Takahashi. ‘Task-Based FMM for Multicore Architectures’. In: *SIAM Journal on Scientific Computing* 36.1 (2014), pp. 66–93. DOI: [10.1137/130915662](https://doi.org/10.1137/130915662). URL: <https://hal.inria.fr/hal-00911856> (cit. on p. 7).
- [39] E. Agullo, S. Cools, E. Fatih-Yetkin, L. Giraud, N. Schenkels and W. Vanroose. ‘On soft errors in the conjugate gradient method: sensitivity and robust numerical detection’. In: *SIAM Journal on Scientific Computing* 42.6 (Nov. 2020). DOI: [10.1137/18M122858X](https://doi.org/10.1137/18M122858X). URL: <https://hal.inria.fr/hal-03022845> (cit. on p. 7).
- [40] E. Agullo, E. Darve, L. Giraud and Y. Harness. ‘Low-Rank Factorizations in Data Sparse Hierarchical Algorithms for Preconditioning Symmetric Positive Definite Matrices’. In: *SIAM Journal on Matrix Analysis and Applications* 39.4 (Oct. 2018), pp. 1701–1725. URL: <https://hal.inria.fr/hal-01940053> (cit. on p. 6).
- [41] E. Agullo, M. Felšöci and G. Sylvand. *A comparison of selected solvers for coupled FEM/BEM linear systems arising from discretization of aeroacoustic problems: literate and reproducible environment*. Technical Report RT-0513. Inria Bordeaux Sud-Ouest, June 2021, p. 100. URL: <https://hal.inria.fr/hal-03263620> (cit. on p. 8).
- [42] E. Agullo, L. Giraud and Y.-F. Jing. ‘Block GMRES method with inexact breakdowns and deflated restarting’. In: *SIAM Journal on Matrix Analysis and Applications* 35.4 (2014), pp. 1625–1651 (cit. on p. 6).
- [43] E. Agullo, L. Giraud and L. Poirel. ‘Robust preconditioners via generalized eigenproblems for hybrid sparse linear solvers’. In: *SIAM Journal on Matrix Analysis and Applications* 40.2 (2019), pp. 417–439. DOI: [10.1137/17M1153765](https://doi.org/10.1137/17M1153765). URL: <https://hal.inria.fr/hal-02074474> (cit. on p. 6).
- [44] P. Blanchard. ‘Fast hierarchical algorithms for the low-rank approximation of matrices, with applications to materials physics, geostatistics and data analysis’. PhD thesis. Bordeaux, 2017. URL: <https://tel.archives-ouvertes.fr/tel-01534930> (cit. on p. 6).

- [45] P. Blanchard, O. Coulaud and E. Darve. *Fast hierarchical algorithms for generating Gaussian random fields*. Research Report 8811. Inria Bordeaux Sud-Ouest, Dec. 2015. URL: <https://hal.inria.fr/hal-01228519> (cit. on pp. 6, 7).
- [46] S. Börm, L. Grasedyck and W. Hackbusch. *Hierarchical Matrices*. Tech. rep. 2003, pp. 1–173 (cit. on pp. 3, 14).
- [47] J. Briant, P. Mycek, M. Destouches, O. Goux, S. Gratton, S. Gürol, E. Simon and A. T. Weaver. ‘A filtered multilevel Monte Carlo method for estimating the expectation of discretized random fields’. working paper or preprint. Nov. 2023. DOI: [10.48550/arXiv.2311.06069](https://doi.org/10.48550/arXiv.2311.06069). URL: <https://hal.science/hal-04337114> (cit. on p. 13).
- [48] A. Buttari, J. Langou, J. Kurzak and J. Dongarra. ‘Parallel tiled QR factorization for multicore architectures’. In: *Concurrency and Computation: Practice and Experience* 20.13 (2008), pp. 1573–1590 (cit. on p. 7).
- [49] E. Carson, N. J. Higham and S. Pranesh. ‘Three-Precision GMRES-Based Iterative Refinement for Least Squares Problems’. In: *SIAM Journal on Scientific Computing* 42.6 (Jan. 2020), A4063–A4083. DOI: [10.1137/20m1316822](https://doi.org/10.1137/20m1316822) (cit. on p. 7).
- [50] F. Casenave, A. Ern and G. Sylvand. ‘Coupled BEM-FEM for the convected Helmholtz equation with non-uniform flow in a bounded domain’. In: *Journal of Computational Physics* 257.A (Jan. 2014). 23 pages, 9 figures, pp. 627–644. DOI: [10.1016/j.jcp.2013.10.016](https://doi.org/10.1016/j.jcp.2013.10.016). URL: <https://hal.archives-ouvertes.fr/hal-00805477> (cit. on p. 8).
- [51] F. Chaouqui, M. J. Gander, P. M. Kumbhar and T. Vanzan. ‘On the nonlinear Dirichlet-Neumann method and preconditioner for Newton’s method’. In: *ArXiv abs/2103.12203* (2021). URL: <https://api.semanticscholar.org/CorpusID:232320731> (cit. on p. 15).
- [52] A. Cichocki, R. Zdunek, A. H. Phan and S.-i. Amari. *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation*. Wiley, 2009 (cit. on p. 6).
- [53] S. Cools, E. F. Yetkin, E. Agullo, L. Giraud and W. Vanroose. ‘Analyzing the Effect of Local Rounding Error Propagation on the Maximal Attainable Accuracy of the Pipelined Conjugate Gradient Method’. In: *SIAM Journal on Matrix Analysis and Applications* 39.1 (Mar. 2018), pp. 426–450. DOI: [10.1137/17M1117872](https://doi.org/10.1137/17M1117872). URL: <https://hal.inria.fr/hal-01753411> (cit. on p. 7).
- [54] O. Coulaud, A. A. Franc and M. Iannacito. *Extension of Correspondence Analysis to multiway datasets through High Order SVD: a geometric framework*. Research Report RR-9429. Inria Bordeaux - Sud-Ouest ; Inrae, Nov. 2021. URL: <https://hal.inria.fr/hal-03418404> (cit. on p. 6).
- [55] V. Dolean, M. J. Gander, W. Kheriji, F. Kwok and R. Masson. ‘Nonlinear Preconditioning: How to Use a Nonlinear Schwarz Method to Precondition Newton’s Method’. In: *SIAM Journal on Scientific Computing* 38.6 (2016), A3357–A3380 (cit. on p. 15).
- [56] A. Falco. ‘Comblent l’écart entre \mathcal{H} -Matrices et méthodes directes creuses pour la résolution de systèmes linéaires de grandes tailles’. Theses. Université de Bordeaux, June 2019. URL: <https://tel.archives-ouvertes.fr/tel-02183902> (cit. on p. 7).
- [57] A. A. Franc, P. Blanchard and O. Coulaud. ‘Nonlinear mapping and distance geometry’. In: *Optimization Letters* 14.2 (2020), pp. 453–467. DOI: [10.1007/s11590-019-01431-y](https://doi.org/10.1007/s11590-019-01431-y). URL: <https://hal.inria.fr/hal-02124882> (cit. on p. 6).
- [58] N. Gillis. *Nonnegative Matrix Factorization*. Society for Industrial and Applied Mathematics, Jan. 2020. DOI: [10.1137/1.9781611976410](https://doi.org/10.1137/1.9781611976410) (cit. on p. 6).
- [59] L. Giraud, Y.-F. Jing and Y. Xiang. *A block minimum residual norm subspace solver for sequences of multiple left and right-hand side linear systems*. Research Report RR-9393. Inria Bordeaux Sud-Ouest, Feb. 2021, p. 60. URL: <https://hal.inria.fr/hal-03146213> (cit. on p. 6).
- [60] L. Grasedyck, W. Hackbusch and B. Nr. ‘An Introduction to Hierarchical (H -) Rank and TT - Rank of Tensors with Examples’. In: *Computational Methods in Applied Mathematics* 11.329 (2011), pp. 291–304 (cit. on p. 3).

- [61] L. Greengard. ‘The rapid evaluation of potential fields in particle systems.’ PhD thesis. Yale University, 1987 (cit. on p. 14).
- [62] B. Gunter and R. Van De Geijn. ‘Parallel out-of-core computation and updating of the QR factorization’. In: *ACM Transactions on Mathematical Software (TOMS)* 31.1 (2005), pp. 60–78 (cit. on p. 7).
- [63] W. Hackbusch. *Hierarchical Matrices: Algorithms and Analysis*. 1st. Springer Publishing Company, Incorporated, 2015 (cit. on pp. 3, 14).
- [64] N. Halko, P.-G. G. Martinsson and J. A. Tropp. ‘Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions’. In: *SIAM Review* 53.2 (2011), pp. 217–288. DOI: [10.1137/090771806](https://doi.org/10.1137/090771806). URL: <http://arxiv.org/abs/0909.4061> (cit. on p. 7).
- [65] T. G. Kolda and B. W. Bader. ‘Tensor Decompositions and Applications’. In: *SIAM Review* 51.3 (Aug. 2009), pp. 455–500. DOI: [10.1137/07070111X](https://doi.org/10.1137/07070111X). URL: <http://epubs.siam.org/doi/abs/10.1137/07070111X> (cit. on p. 3).
- [66] C. Kruse, V. Darrigrand, N. Tardieu, M. Arioli and U. Rude. ‘Application of an iterative Golub-Kahan algorithm to structural mechanics problems with multi-point constraints’. In: *Adv. Model. Simul. Eng. Sci.* 7.1 (2020), p. 45. DOI: [10.1186/s40323-020-00181-2](https://doi.org/10.1186/s40323-020-00181-2). URL: <https://doi.org/10.1186/s40323-020-00181-2> (cit. on p. 6).
- [67] C. Kruse, M. Sosonkina, M. Arioli, N. Tardieu and U. Rude. ‘Parallel solution of saddle point systems with nested iterative solvers based on the Golub-Kahan Bidiagonalization’. In: *Concurr. Comput. Pract. Exp.* 33.11 (2021). DOI: [10.1002/cpe.5914](https://doi.org/10.1002/cpe.5914). URL: <https://doi.org/10.1002/cpe.5914> (cit. on p. 6).
- [68] V. Le Bris, M. Odunlami, D. Begue, I. Baraille and O. Coulaud. ‘Using computed infrared intensities for the reduction of vibrational configuration interaction bases’. In: *Phys. Chem. Chem. Phys.* 22 (13 2020), pp. 7021–7030. DOI: [10.1039/D0CP00593B](https://doi.org/10.1039/D0CP00593B). URL: <http://dx.doi.org/10.1039/D0CP00593B> (cit. on p. 6).
- [69] B. Lize. ‘Resolution directe rapide pour les lements finis de frontire en lectromagntisme et acoustique : \mathcal{H} -Matrices. Paralllisme et applications industrielles’. Theses. Universit Paris-Nord - Paris XIII, June 2014. URL: <https://tel.archives-ouvertes.fr/tel-01244260> (cit. on p. 7).
- [70] P.-G. Martinsson and J. Tropp. ‘Randomized Numerical Linear Algebra: Foundations & Algorithms’. In: (2020). eprint: [2002.01387](https://arxiv.org/abs/2002.01387). URL: <http://arxiv.org/abs/2002.01387> (cit. on p. 7).
- [71] T. Mary. ‘Block Low-Rank multifrontal solvers: complexity, performance, and scalability’. PhD thesis. Universit Paul Sabatier-Toulouse III, 2017 (cit. on p. 14).
- [72] M. Odunlami, V. Le Bris, D. Begue, I. Baraille and O. Coulaud. ‘A-VCI: A flexible method to efficiently compute vibrational spectra’. In: *The Journal of Chemical Physics* 146.21 (June 2017), p. 214108. DOI: [10.1063/1.4984266](https://doi.org/10.1063/1.4984266). URL: <http://aip.scitation.org/doi/10.1063/1.4984266> (cit. on p. 6).
- [73] I. V. Oseledets. ‘Tensor-Train Decomposition’. In: *SIAM Journal on Scientific Computing* 33.5 (Jan. 2011), pp. 2295–2317. DOI: [10.1137/090752286](https://doi.org/10.1137/090752286). URL: <https://doi.org/10.1137/090752286> (cit. on p. 3).
- [74] L. Poirel. ‘Algebraic domain decomposition methods for hybrid (iterative/direct) solvers’. Theses. Universit de Bordeaux, Nov. 2018. URL: <https://tel.archives-ouvertes.fr/tel-02070618> (cit. on p. 6).
- [75] J.-R. Poirier, O. Coulaud and O. Kaya. ‘Fast BEM Solution for 2-D Scattering Problems Using Quantized Tensor-Train Format’. In: *IEEE Transactions on Magnetism* 56.3 (Mar. 2020), pp. 1–4. DOI: [10.1109/TMAG.2019.2954584](https://doi.org/10.1109/TMAG.2019.2954584). URL: <https://hal.inria.fr/hal-03150956> (cit. on p. 7).
- [76] G. Sylvand. ‘La mthode multiple rapide en lectromagntisme. Performances, paralllisation, applications’. Theses. Ecole des Ponts ParisTech, June 2002. URL: <https://pastel.archives-ouvertes.fr/tel-00005683> (cit. on p. 7).
- [77] N. Venkovic, P. Mycek, L. Giraud and O. Le Maitre. *Recycling Krylov subspace strategies for sequences of sampled stochastic elliptic equations*. Research Report RR-9425. Inria Bordeaux - Sud Ouest, Oct. 2021. URL: <https://hal.archives-ouvertes.fr/hal-03366966> (cit. on p. 6).