PEPR EXPLORATOIRES
PROJET CIBLÉ
2022

PROGRAMME NUMPEX

DOCUMENT PRÉSENTATION PROJET

Exa-MA

Acronyme	**Exa-MA**		
Titre du projet en français	*Méthodes et Algorithmes à l'Exascale*		
Titre du projet en anglais	*Methods and Algorithms for Exascale*		
Mots-clefs	Calcul Exascale, méthodes, algorithmes, Discrétisation, Réduction d'ordre, IA, Solveurs, Problèmes inverses et assimilation de forme, Optimisation, Quantification d'Incertitudes, Benchmarks, Logiciels		
Établissement porteur	Université de Strasbourg		
Responsable du projet	**Prénom, Nom, Qualité**		
	Christophe Prud'homme, PR		
	Courriel	**Téléphone**	
	prudhomme@unistra.fr	**0687646051**	
Durée du projet	82 Mois		
Aide totale demandée	6 254 997,79 €	**Coût complet**	24 417 823, 28 €

Liste des établissements du consortium :

Établissements d'enseignement supérieur et de recherche	***Secteur(s) d'activité***
Université de Strasbourg	
Sorbonne Université	
École Polytechnique	
Bordeaux INP *ENS Lyon* *Paris Pantheon Assas* *Université Côte d'Azur* *Université de Lorraine* *Université Gustave Eiffel* *Université de Lille* *Université de Pau et des Pays de l'Adour* *Université de Picardie*	

Organismes de recherche	***Secteur(s) d'activité***
CEA	
INRIA	
CNRS	

Résumé du projet en français (Non Confidentiel – 4000 caractères maximum, espaces inclus)

Il existe un nombre croissant de problèmes pour lesquels les expériences sont impossibles, dangereuses ou extrêmement coûteuses. Le calcul à l'échelle extrême permet de résoudre des modèles prédictifs beaucoup plus précis et d'analyser des quantités massives de données grâce à l'IA. La combinaison de la modélisation prédictive avec les données, associée à des stratégies d'apprentissage automatique et d'IA, peut créer de nouvelles opportunités dans le domaine scientifique. En particulier, il s'agit de passer de *"l'humain dans la boucle"* à une conception, une découverte ou une évaluation hybride pilotée par l'humain et l'intelligence artificielle. Cependant, divers défis scientifiques et techniques doivent être relevés pour exploiter les capacités du calcul exascale. Ces goulets d'étranglement ont un impact profond sur les méthodes et les algorithmes dans tous les aspects de la chaîne d'outils de simulation par *(i)* l'évitement de la communication, *(ii)* l'adaptation du grain parallèle et de l'intensité de calcul au niveau des nœuds, *(iii)* la gestion du matériel hétérogène et des représentations de données et *(iv)* l'auto-paramétrisation.

Le projet Exa-MA se concentre sur les aspects Exascale des méthodes numériques, en assurant leur évolutivité vers le matériel existant et à venir. En outre, il s'agit d'un projet transversal, proposant des méthodes et des outils où la modélisation, les données et l'IA, par le biais des algorithmes, sont centrales.

Résumé du projet en anglais (Non Confidentiel – 4000 caractères maximum, espaces inclus)

There is a growing number of problems where experiments are impossible, hazardous, or extremely expensive. Extreme-scale computing enables the solution of vastly more accurate predictive models and the analysis of massive quantities of data thanks to AI. Combining predictive modeling with data, coupled with machine learning and AI strategies, can create new opportunities in science. In particular, move from Human-in-the-Loop towards hybrid Human and Artificial Intelligence-driven design, discovery, or evaluation. However, various scientific and technical challenges need to be met to exploit exascale computing capabilities. These bottlenecks impact methods and algorithms in a profound way on all aspects of the simulation toolchain through (i) avoidance of communication, (ii) adaptive parallel grain and more compute-intensive at node level, (iii) handling of heterogeneous hardware and data representations and (iv) self-parametrization.

The Exa-MA project concentrates on the Exascale aspects of the numerical methods, ensuring their scalability to existing and forthcoming hardware. Furthermore, it is a transverse project, proposing methods and tools where modeling, data and AI, through algorithms, are central.

Table des matières/Table of Contents

1. Context, objectives and previous achievements

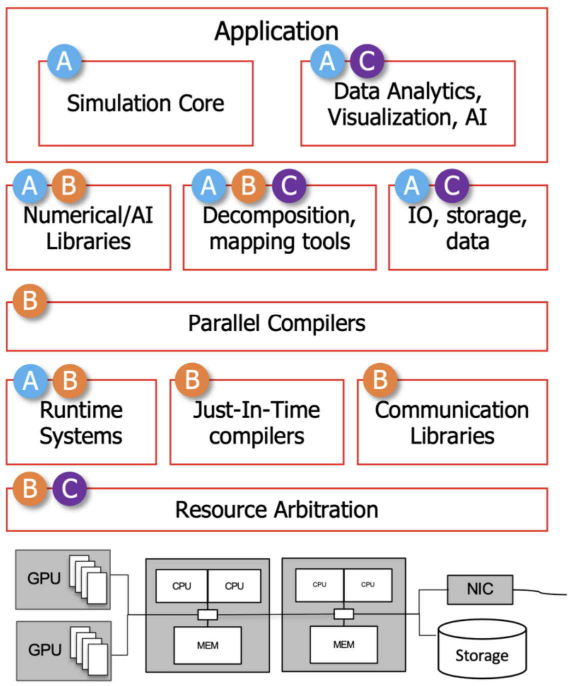
1.1. Context, objectives and innovative features of the project

There is a growing number of problems where experiments are impossible, hazardous, or extremely expensive. Extreme-scale computing enables the solution of vastly more accurate predictive models and the analysis of massive quantities of data. More than that, it enables the possibility to create a digital copy of a physical asset that can be fed with data to understand, improve or fix the latter. These challenges include: ***(C1)*** Reduce carbon footprint (GHG) in transportation, buildings and cities; ***(C2)*** Design, control, and manufacture of advanced materials; ***(C3)*** Understand and simulate the human brain; ***(C4)*** Understand fission and fusion reactions and design advanced experiment facilities for fusion; ***(C5)*** Monitor the health of our planet (climate prediction, impact assessment of environmental policies, etc.); ***(C6)*** Monitor and personalize the health of human beings; ***(C7)*** Design better drugs; ***(C8)*** Design cost-effective renewable energy resources (batteries, biofuels, solar photovoltaics, etc.); or more generally ***(C9)*** Understand the Universe. These challenges require tremendous computing power to understand them and help decision makers.
Exascale computing is the next frontier to unlock new discoveries. We face, however, new bottlenecks as we reach these computing facilities including ***(B1)*** **energy efficiency**: develop energy efficient technologies to meet the at most 20 MW target. ***(B2)*** **interconnect technology**: improve vertical (intra-node) and horizontal (inter-node) data movement in terms of energy efficiency and performance. ***(B3)*** **Memory technology**: integrate new memory technologies (e.g., PCRAM, NOR Flash, ReRAM, memristor) to improve capacity, bandwidth, resiliency, and energy efficiency. ***(B4*****)** **Scalable system software**: Increase the scalability, power sensitivity, and resiliency of system software (e.g., operating systems, runtime systems, monitoring systems). **(B5) Programming systems**: develop new programming paradigms to express fine-grained concurrency, locality, and resilience. ***(B6)*** **Data Management**: develop software that can handle massive amounts of data–this concerns both offensive I/O (e.g., data analysis & compression) and defensive I/O (e.g., fault tolerance). ***(B7)*** **Exascale Algorithms**: redesign algorithms to improve scalability (e.g., reduce communication, avoid/hide synchronization) and computational efficiency on accelerators. ***(B8)*** **Discovery, design, and decision algorithms**: Research should focus not only on "single heroic simulations" but also on ensembles of many small runs (e.g., common for uncertainty quantification or parameter optimization). ***(B9)*** **Resilience, robustness and accuracy**: Computations must be correct, reproducible and verifiable, even in the presence of software and hardware errors (hard and/or soft error). ***(B10)*** **Scientific productivity**: scientists must have the tools to use exascale systems productively (e.g., develop programs, run applications, prepare inputs, collect outputs, analyze results). ***(B11)*** **Reproducibility, replicability of computation**: reproducibility is an essential ingredient of the scientific enterprise. The ability to reproduce results builds trust so that we can rely on the results as foundations for future scientific exploration. Presently, the fields of computational and computing sciences provide two opposing definitions of *reproducible* and *replicable*. In computational sciences, reproducible research means authors provide all necessary data and computer codes to run analyses again, so others can re-obtain the results. The concept was adopted and extended by several communities, where it was distinguished from replication: collecting new data to address the same question, and arriving at consistent findings. ***(B12)*** **Pre/Post processing**: visualization, in situ processing. ***(B13)*** **Opportunity to integrate uncertainties directly into the core of the calculation** (unseen).

If the bottlenecks *(B1-B6;B12)* are to be tackled at the methods and algorithms level through transverse collaborations within the PEPR NumPEx, Exa-MA will directly address *(B7-B11;B13)* and thus its **main objectives** are ***(O1)*** **to develop methods, algorithms, and implementations that, taking advantage of the exascale architectures,** empower modeling, solving, assimilating model and data, optimizing and quantifying uncertainty, at levels that are unreachable at present; ***(O2)*** **to develop or contribute to software libraries** allowing to assemble specific critical reusable components, hiding the hardware complexity and exposing only the specific methodological interface; ***(O3}*** **to identify and co-design Methodological and Algorithmic Patterns** at exascale that can be reused efficiently in large scale applications (e.g., in weather forecast); ***(O4)*** **to enable AI algorithms** to attain performances at exascale, exploiting the methods ***(O1)*** and the libraries ***(O2)*** developed; and ***(O5)*** **to provide demonstrators** through mini-apps and proxy-apps that will be openly available and benchmarked.
Our project objectives mirror the methodology components of the Exascale Computing Project (USA) as well as the Fugaku initiative in Japan.

1.2. Main previous achievements

The French community has many achievements over the last decades in High-Performance Computing, from theoretical breakthroughs to widely-used implementations including in the private sector, benchmarking, and diffusion in many application domains. **All participants of Exa-MA, including partners, have made such significant contributions.** To name a few: *(i)* in discretization mesh generation, validation and adaptation – cGal and Mmg/parMmg[CGAL/MMG] – high order and spectral element methods, non conforming methods such as HDG, HHO [Cockburn2016] or Mortar [Bernardi2005] or parallel in time [Maday2001] methods *(ii)* reduced order methods (Reduced basis [Buffa2010] and empirical interpolation [Barrault2004]) and surrogate methods *(iii)* provably robust scalable domain decomposition methods (such as GenEO [Spillane2014]) and associated framework (such as HPDDM [Jolivet2014] *(iv)* on data assimilation methods [Asch2016] *(v)* Optimization (heuristics for large-scale optimization and shape optimization) and *(vi)* global sensitivity analysis and uncertainty quantification [Iooss2015].
These contributions remain central to Exa-MA as they embed core enabling properties for exascale HPC and mathematical correctness. They need to be revisited and extended to reach our exascale objectives ***(O1-O5)***, and possibly to be enhanced and combined with AI methods to reach exascale performance.

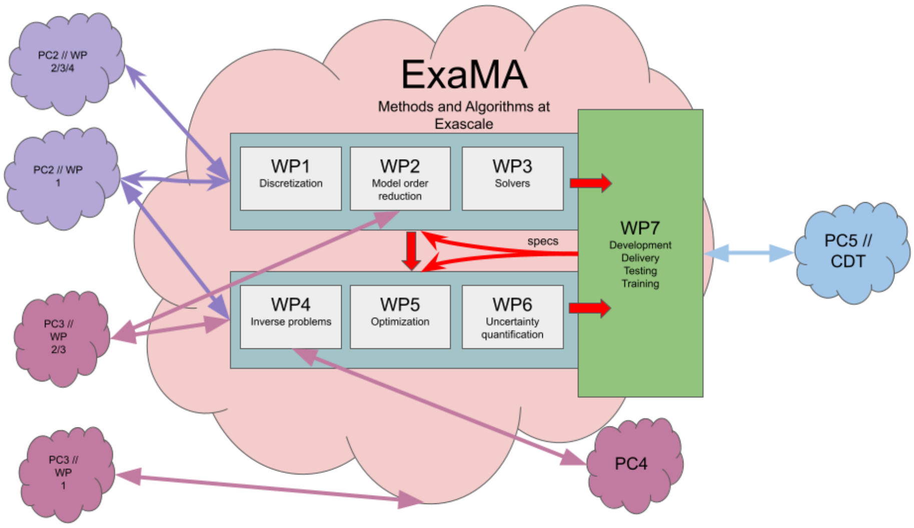


2. Detailed project description

2.1. Project outline, scientific strategy

Exa-MA (A) is one of the focus projects of NumPEx to reach Exascale–the figure on the right positions the

project alongside the other PEPR focus projects ExaSoft (B) and ExaDoST (C).
Various scientific and technical challenges ***(B1-B13)*** need to be met to exploit exascale computing capabilities. Methods and algorithms must be revisited in a profound way on all aspects of the simulation tool chain and exploit *(i)* reduction of communication, *(ii)* adaptive parallel grain and more compute-intensive at node level, *(iii)* handling of heterogeneous hardware and data representations, and *(iv)* self parametrization. Moreover, combining predictive modeling with data, coupled with machine learning and AI strategies, can create new opportunities in science as in particular, moving from Human-in-the-Loop towards hybrid Human and Artificial Intelligence-driven design, discovery, or evaluation. Exa-MA's vision is to combine these ingredients to reach and go beyond the Exascale barrier.
To this end, Exa-MA funds inter-partner research groups–and sometimes along with external partners—on selected topics in 6 Work Packages (WP1-WP6), the last WP(WP7) is transverse and is the place for development, delivery, testing, training and possibly deployment together with the other PEPR projects.



The figure illustrates the scientific strategy of Exa-MA with 3 core WPs—WP1 discretization, WP2 model order reduction, WP3 solvers—and, building on the core WPs, 3 advanced WPs—WP4 inverse problems, WP5 optimization, and WP6 uncertainty quantification.

2.2. Scientific and technical description of the project

The main objective is to scale-up methods and algorithms in predictive simulation data analysis, up to digital twinning, including uncertainty quantification and inverse problems. The work package distribution follows the research topics. We indicate important links to other PEPR focused projects' work packages as well as the addressed exascale bottlenecks, and we present the various tasks numbered in boldface. Each work package includes deliverables comprising: *(i)* reports that yield articles and conferences, *(ii)* benchmarking to demonstrate the extent of the mitigation of the corresponding exascale bottlenecks, and *(iii)* software developments.
The starting date of the NumPEx program is the 1st of January 2023, and the scientific starting date will be the 14th of April 2023. The tasks have been scheduled for 60 months. However, to anticipate potential delays (e.g. PhD or postdoc hire), the duration is set to 82 months.

2.3. Planning, KPI and milestones

The scientific program of Exa-MA is divided into 7 work packages connected to each other as described in the Figure above.

WP Number		WP1	Lead	CEA				Start	1	End	60
WP Title		Discretization									
Part. #	1	2	3	4	5						
Name	UNISTRA	CEA	INRIA	EP	SU						
PMs	29,4	96	129	66	3						

Objectives

The key objectives of this work package are twofold: *(i)* **Geometric domain representations and their discrete counterparts (such as meshes) [B2, B6, B7, B9, B11-B13]** are usually the main starting point and support of the simulation. These include adaptive, possibly multiresolution, robust to defects, resilient to heterogeneous input, and efficient parallel representations of large-scale models; *(ii)* **Physics-based models [B7, B10]** need to include multiple phenomena or process couplings at multiple scales in space and time. Space and time adaptivity are then mandatory. Time integration requires special care to become more parallel, more asynchronous, and more accurate for long-time simulations. The requirement of adaptivity in geometry and physics creates an imbalance that needs to be overcome. In the discretization process, we will *(a)* favor high-order methods to increase the computational intensity and reduce communications and *(b)* favor nonconforming methods that are designed to avoid/reduce/minimize communications

Description of work

T1.1: Mesh generation

T1.1.1 Valid large-scale mesh generation. We will explore the automated generation of large-scale meshes, either from measurement data or from designed CAD models. Such input data require 3D reconstruction and meshing algorithms capable of dealing with heterogeneous, defect-laden representations. We will look for meshing algorithms that scale, and are unconditionally robust to input defects. Valid meshing herein translates into output meshes that are watertight and without self-intersections. The meshing algorithms will also provide detail levels with fine-grain and adjustable balance between complexity, resolution and approximation.

T1.1.2 Mesh generation for non-conforming methods. In the framework of non-conforming domain decomposition, two-level mesh generation algorithms will be explored and implemented. The macro elements are independently refined, in parallel, into fine meshes with no conformity requirement along the boundaries of the macro elements. The coarse grid must support partitioning and dynamic load balancing when proceeding to fine mesh adaptation. Mesh connectivity must be available and shared efficiently between neighboring macro-elements.

T1.1.3 Mesh generation of all-hexahedral block grids. In this task, we develop robust hexahedral block structured meshes, based on state-of-the-art algorithms and open-source software to derive linear block structure that will be a support to: (1) Curve them to yield high-order blocks, and (2) Define local grid patterns that fit the multigrid method requirements and can be applied in each block or even onto several connected blocks.

T1.2 Adaptive Mesh Refinement for unstructured grids. Mmg/ParMmg is a framework that can adapt to very large meshes, currently up to several billion cells. ParMmg is a MPI-based component for parallel mesh adaptation. Our objective is to break new scalability barriers in Mmg/ParMmg through *(i)* code

Description of work

robustification, *(ii)* improved memory management, *(iii)* graph coloring and interface migration with load balancing, *(iv)* metrics management, *(v)* benchmark, profile and optimization.

T1.3 Adaptive Mesh Refinement for Cartesian or block grids. We use Adaptive Mesh Refinement (AMR) methods to improve accuracy and computation time in simulations, particularly when dealing with spatially localized physical phenomena. Several strategies for error control are presented, including patch-based, cell-based and adaptive multiresolution using wavelets. Large-scale simulations using AMR typically rely on specialized data-structures. We focus on several discretization methods (finite elements, finite volumes, and finite differences), conforming or non-conforming, with structured mesh refinement using all-hexahedral/quadrilateral elements for exascale simulations. We also propose to work on HPC algorithm extensions of the local multigrid AMR methods to develop load-balancing algorithms and efficient refinement strategies.

T1.4 Finite Element Exascale Framework (FEEF) We utilize exascale hardware and low-level software infrastructure to create a framework for high-order/spectral finite element methods on unstructured grids. We aim at providing a framework that can handle the full de Rahm complex, while hiding low-level architecture details. The focus is on non-conforming methods to provide optimized kernels for finite elements, which will require close collaboration with PC2's WP2 and WP3.

T1.5 Exploit non conforming methods for efficient parallel parallel Conformal finite-element discretizations lead to large sparse linear systems that are difficult to solve with modern parallel architectures. We propose to explore non-conformal methods such as Trefftz methods, hybridizable discontinuous Galerkin (HDG) methods, including HHO and the mortar element method, as alternatives to cG. We aim at improving the conditioning of the resulting system for efficient iterative solution procedure or reducing the size of the sparse linear system to be solved. Our plan is also to work on the combination of these methods and domain decomposition methods in order to accelerate the convergence of iterative solvers.

T1.6 Time-strategy for evolution equations when the mesh is dynamically adapted

T1.6.1 Parallel in time strategies: improve coarse solver. The time-parallelization methods, such as the parareal method, propose to use domain decomposition techniques in the time direction to solve more rapidly on distributed architectures. However, there are still issues to be addressed to improve efficiency, such as the efficiency of the coarse solve, the coupling of different PDEs with at least a parabolic PDE and the extension to Hamiltonian systems.

T1.6.2 Error control in time. We consider the challenges of solving evolving multiscale problems. Two types of multiscale problems are mentioned: those with relaxation sources and short time scales, and those with complex sources involving a broad spectrum of scales. The first type can lead to inefficient use of mesh adaptation and strong stability constraints must be met. Two key issues are addressed: designing numerical schemes with strong stability properties and preserving invariant convex spaces. For the second situation, some schemes based on adaptive operator splitting and adaptive multiresolution have been developed. Three breakthroughs are needed to tackle exascale simulations: quantifying compression errors, extending to high order methods and wavelets, and designing new time integration strategies.

T1.7 Efficient multimodel/multiphysics coupling

T1.7.1 Multiphysics partitioned discretization. We need far more realistic simulations of multiphysics problems and the importance of scalability in these simulations. Our focus is on partitioned coupling, using well-established discretization methods for each physics that have demonstrated good scalability properties. The task involves studying scalable coupling algorithms and ensuring that the properties of the system obtained after discretization meet the scaling requirements, with a focus on fields' projections and conservations. This task is related to task T3.5 of WP3 on exascale simulations of partitioned multiphysics coupling. We will also consider adaptive high order coupling techniques to take advantage of partitioning for high performance computing without compromising on accuracy of the system, even when dealing with DAEs and elliptic/parabolic equations.

Description of work
T1.7.2 Multiscale coupling Numerical methods, such as discontinuous Galerkin discretizations, are conducive for exascale computing because they allow for non-conformal discretization. These methods can rely on a primal formulation set on the skeleton of a coarse mesh and a dual one involving the solution of small problems solved at the element level. The use of different basis functions and numerical methods, such as those based on multiscale finite elements or neural networks, can be used to reproduce different physics.

Deliverables	Title	Tasks	Dates
D1.1-S	Software toolboxes for mesh generation and adaptation , space-time discretisation and coupling	T1.*,	M24,M36,M48,M60
D1.2-MR-RP-P	Activity report included in annual report D0.2-TR	T1.*	M12,M24,M36,M48,M60
D1.3-B	Benchmarking analysis report, including bottlenecks and breakthroughs	T1.*, T7.1	M12,M24,M36,M48,M60

Collaborations with Projects: ExaSoft

WP Number		WP2	Lead		INRIA			Start	1	End 60
WP Title		Model order, Surrogate, Scientific Machine Learning methods								
Part. #	1	2	3	4	5					
Name	UNISTRA	CEA	INRIA	EP	SU					
PMs	173,4	0	114	27	2,4					

Objectives
We will use *non-intrusive* approaches for designing ultra-fast surrogate models of complex physical problems, and strategies for leveraging these surrogates. We will consider **data-driven** techniques including comparisons between reduced basis methods and NN-based methods and **model-driven** techniques, with a focus on physics-based NN models.

Description of work
T2.1 Surrogate models based on physics-driven Deep Learning Artificial Neural Networks (ANNs) frameworks that do not require labeled data sets for their training have been extensively studied in recent years. The concept of Physics-Informed Neural Networks (PINNs) is currently the most popular and is used to solve PDEs, fractional equations, integral-differential equations, and stochastic PDEs. They have a great potential of application to the solution of non-linear, time-dependent, parameter dependent PDEs or even to inverse problems. However, PINNs struggle to scale to problems with larger domains and more complex, multi-scale solutions, partly due to the spectral bias of NNs and the increasing size of the underlying PINN optimization function. We propose to combine PINNs with domain decomposition methods (DDM). Specific topics are: training data, network architecture, loss definition and optimization

and finally applications.

T2.2 PDE operator learning with Neural Operators Neural Operators (NO) are neural networks specifically dedicated to the approximation of inverse PDE operators. These methods, such as DeepOnet and FNO, allow for fast computation of PDE solutions for various physical problems. The research aims to study NO networks theoretically and methodologically, focusing on applying them to more complex nonlinear problems and difficult geometries, and working on ways to guarantee the quality of results.

T2.3 Data-driven model order reduction. Recent approaches include building nonlinear reduction methods, such as using auto encoders and projections or nonlinear corrections to classical methods. This task aims to study theoretical aspects and propose new methods for nonlinear reduction on difficult geometries, hyper-reduction, closure, and learning techniques for multiple objective cost functions, and also to investigate stochastic versions with deep generative latent models. These methods should be able to leverage exascale computing systems.

T2.4 Non-intrusive reduced basis methods for parametric problems In this task, we study Non-Intrusive Reduced Basis (NIRB) methods to reduce the computational cost of high-fidelity codes. These methods include operator inference methods that enable the automated affine or higher-order decomposition of operators and two-grid methods. The offline part of these methods is expensive but performed only once, while the online phase can be vastly accelerated by orders of magnitude. Both steady and evolution problems are considered and we also combine these strategies with parallel in time methods for parabolic problems.

T2.5 Mixing low- and high-fidelity models Multi-fidelity modeling (MFM) [Peherstorfer2018] uses a hierarchy of lower-fidelity models, or surrogate models, to reduce the computational cost of repeatedly using a high-fidelity model. The lower-fidelity models must, in general, be chosen for each problem, but there are some general principles that can be followed that we will focus on. Scaling is, once again, a central issue and argues in favor of the Monte Carlo type approaches [Asch2022] that we will consider.

T2.6 Real-time models with super resolution methods allow computing an *interpolate,* which is a fine representation of pictures or functions using a coarse representation. Deep Learning methods provide impressive results in this direction [Saharia2022] with a nice perspective for PDEs [Hao2021]. Here we propose to extend these methods to complex PDEs but also to stochastic processes (using deep, generative models) and on unstructured grids (using graph neural networks). These approaches would allow very low-resolution simulations to be carried out and corrected in real time.

Deliverables	Title	Tasks	Dates
D2.1-MR-TR-P	Activity report included in annual report D0.2-TR	T2.*	M12, M24,M36,M48,M60
D2.2-S	Surrogate modeling toolbox	T2.1,T2.2,T2.3,T2.4	M24,M36,M48,M60
D2.3-BR	Benchmarking analysis report, including bottlenecks and breakthroughs	T2.*, T7.1	M12, M24, M36, M48, M60

Collaborations with Projects: ExaSoft

WP Number	WP3	Lead	INRIA	Start	1	End	60

WP Title	Solvers for linear algebra and multiphysics									
Part. #	1	2	3	4	5					
Name	UNISTRA	CEA	INRIA	EP	Sorbonne Université					
PMs	12	36	219	9	150					

Objectives
We first intend to design novel or improved numerical kernels that are mostly agnostic of the underlying models (e.g., PDE) and approximation techniques (e.g., FEM) for the solution of linear systems. The envisioned solution techniques are generic in the sense that they do not depend on the way the function f is represented (e.g., matrix or implicit access to f in the linear case). To reduce the computational complexity, memory footprint and data movement, techniques such as communication avoiding/hiding, mixed arithmetic and data compression (e.g., low-rank approximations) will be exploited. The second objective is to fully exploit the extreme parallelism enabled by the forthcoming platforms to design coupled physic solvers that rely on state-of-the-art optimized mono-physics solvers.

Description of work

T3.1 *Domain decomposition methods with subspace-correction* To enhance the scalability of multilevel domain decomposition methods, we will investigate theoretically and experiment their robustness with respect to inexact setups and applications of the preconditioner that might arise for instance from mixed arithmetic calculation, inexact local solves, low rank approximations. The validation of these studies will be conducted primarily in the HPDDM library but could also benefit others such as Maphys++.

T3.2 Exploiting data-sparsity, multiple precision and data compression

T3.2.1 Modular, composable mixed precision Krylov solvers. Subspace Krylov exhibits appealing features enabling the use of mixed precision arithmetic in different computational steps that might be computed by different numerical variants in a modular manner. To guarantee the final numerical quality of the computed solution novel modular, analyses must be developed for enabling composable parallel implementations of these subspace solvers (in close collaboration with PC2).

T3.2.2 Decoupling the data representation from the arithmetic: the variable accuracy paradigm
The variable accuracy paradigm is a promising avenue to control the memory footprint and the volume of communication by decoupling the data representation from the arithmetic, while ensuring a user-prescribed accuracy. Some analysis must be developed to guarantee the numerical quality of the computed solutions; this will enable robust implementations of the numerical solvers (in close collaborations with PC2) fully exploiting the underlying features (memory and processing units) of modern computing architectures.

T3.2.3 Precision auto-tuning tools Precision auto-tuning tools provide mixed precision versions of numerical codes, taking into account accuracy requirements on the results. New algorithms must be proposed to improve precision auto-tuning performance and extend it to arbitrary precision. Furthermore, new methodologies must be developed to perform autotuning of both numerical formats and performance parameters in coupled physics simulations as addressed in Task 3.3.2.

T.3.2.4 Silent errors in solution techniques The objective of this task is to study and design silent error detection and hopefully correction that might appear at scale. These evaluations will either be tailored to the intrinsic properties of the numerical schemes or based on statistical techniques.

T3.3 Adaptive solution strategies for exascale multiphysical and multiscale models

T3.3.1 Many multiphysics problems may be recast as saddle point problems The task will consist of investigating the parallel efficiency of domain decomposition methods for the solution of saddle point problems for coupled multiphysics problems.

T3.3.2 Exascale resolution for simulations in partitioned coupling The task will consist of the automatic tuning of performance parameters appearing in the partitioned coupling between exascale-ready software components. Such parameters (coupling strengths, path between physics, adaptive internal convergence criteria…) are intended to be handled through a joint strategy with the auto-tuning work from Task 3.2.3.

Deliverables	Title	Tasks	Dates
D3.1-MR-TR-P	Activity report included in annual report D0.2-TR	T3.*	M12,M24,M36,M48,M60
D3.2-S	Software packages HPDDM, Maphys++, PROMISE	T3.1,T3.2.*,T3.3 *	M24,M36,M48,M60
D3.3-BR	Benchmarking analysis report, including bottlenecks and breakthroughs	T3.*,T7.1	M12,M24,M36,M48,M60

Collaborations with Projects: ExaSoft

WP Number		WP4	Lead		UNISTRA		Start	1	End	60
WP Title		Combine Data and Models, Inverse Problems at Exascale								
Part. #	1	2	3	4	5					
Name	UNISTRA	CEA	INRIA	EP	SU					
PMs	19,2	0	111	0	0					

Objectives

The combination of data and models will be achieved by formulating and solving inverse problems. The approach will cover both deterministic and stochastic models and methods, and will consider observation sparsity and sufficiency. Multi-fidelity models (MFM), including reduced-order and AI-based surrogates, will be studied for the solution of inverse problems. The objectives of the WP4 are to improve existing deterministic methods and their scaling to exascale, formulate new stochastic methods for inverse problems, improve observation strategies and implement multi-fidelity schedules at exascale.

Description of work

T4.1 Deterministic methods Variational methods are inherently sequential, so additional sources of parallelism are needed. One promising avenue is the use of time-parallel methods with the weak-constraint formulation of variational assimilation. Another is exploiting the different physics, spatial and temporal scales of the modules of a modeling system to improve parallelism. Access to exascale computing resources also allows further development of hybrid ensemble-variational approaches, where an ensemble is used for better definition of error covariance.

T4.2 Stochastic methods The most realistic representation of any physical system can be obtained by using stochastic processes obeying stochastic differential equations. Until now, this has largely been forgone due to the theoretical and algorithmic complexity and the resulting computational costs. But exascale systems now provide the resources to reconsider these approaches.Two approaches can be considered. The first is based on Itô theory and is formulated in terms of stochastic (ordinary) differential equations (SDE). The second uses stochastic partial differential equations (SPDE). This task will study and compare the above approaches on several classical inverse and data assimilation problems.

T4.3 Observations The description of observation error statistics in data assimilation is important but difficult. It has traditionally been formalized by assuming uncorrelated Gaussian errors. However, this assumption is not always valid, and our goal is to get rid of it albeit the significant computational cost alleviated by the exascale resources. Another challenge is the forecast errors dominated by position errors, they are poorly captured by classical metrics. We need alternative metrics based on transport theory.

T4.4 Taking advantage of multi-fidelity modeling This task will be treated conjointly with WP2, Task 2.5. We will consider PINN-type models as potential candidates for MFMs—see Tasks 2.1 and 2.2 of WP2.

T4.5 Challenges of multi-fidelity in inverse problems: criteria to update reduced models As explained in T4.4, there is a passage from model-to-model in the overall multi-fidelity hierarchy. Thus, MFM needs a model management or *scheduling strategy* to combine high-fidelity and lower-fidelity models in some optimal way, either to reduce the cost, or to maximize the precision. In the stochastic framework, where we are interested in uncertainty quantification, we can seek minimal variance, for example. In this task, we will take up the models developed in Task 4.4 and develop pertinent scheduling strategies. The overall objective is to provide a choice of strategies that can then be used for different problems in different contexts. The recourse to exascale scaling will be a central issue here.

Deliverables	Title	Tasks	Dates
D4.1-MR-TR-P	Activity report included in annual report D0.2-TR	T4.*	M12,M24, M36, M48, M60
D4.2-S	Software package for inverse problems and data assimilation	T4.1, T4.2, T4.3	M24, M36, M48,M60
D4.3-BR	Benchmarking analysis report, including bottlenecks and breakthroughs	T4.*, T7.1	M12,M24,M36,M48,M60

Collaborations with Projects: ExaDoST, ExaAToW

WP Number		WP5	Lead		INRIA			Start	1	End	60
WP Title		Optimization									
Part. #	1	2	3	4	5						
Name	UNISTRA	CEA	INRIA	EP	SU						
PMs	53,4	0	189	0	0						

Objectives

We focus on the design and implementation of exascale optimization algorithms for solving large-scale optimization problems. The main challenge is managing a large amount of irregular tasks on supercomputers with multiple levels of parallelism and heterogeneous resources. The main objectives include: (i) Exascale combinatorial, continuous and mixed optimization using exact and approximate algorithms; (ii) Exascale surrogate-based optimization using surrogates, multi-fidelity, optimizers, and their coupling; (iii) Exascale shape optimization, specifically when involving multiphysics models. **Bottlenecks**: [B7, B9, B10, B13]

Description of work

T5.1- Exascale combinatorial and continuous optimization: this task concerns the design of general exascale optimization algorithms. Both exact (e.g. branch and bound, tree search) and iterative approximate algorithms (e.g., evolutionary algorithms, swarm intelligence) are considered. To assess the performances of the designed algorithms, we will consider standard benchmarks in combinatorial and continuous optimization. We will consider decomposition-based exascale optimization. The motivation of decomposition is twofold: (1) tackling LOPs that are intractable using traditional algorithms; (2) decomposition is a major step in the parallelization methodology. The challenge is to define new efficient decomposition strategies in the decision and objective spaces to design ultra-scalable optimization algorithms.
T5.2 - Exascale surrogate-based optimization: SBO algorithms have to be adapted to exploit exascale HPC systems, evaluate multiple candidate solutions in each iteration, and exploit multi-fidelity models. In learning-based exascale optimization, ML is mainly used to raise the challenge of expensive functions using surrogate-based optimization (SBO) and multi-fidelity models (MFMs). We will design new parallel SBO algorithms. Three dimensions will be considered: the surrogate, the optimizer and its sampling strategy, and the coupling between them. MFMs will enable acceleration of seeking optima.
T5.3 - Exascale shape optimization: we develop a toolbox for shape optimization for exascale. We consider several directions to improve the performance of shape optimization algorithms used in general. We plan to study through theoretical analysis and numerical tests the possibility of considering different meshes for the state variable and the design variable. Furthermore, we wish to use reduced models in WP2 involving for example neural networks, allowing faster evaluations of the models (state and adjoint) at each iteration of the algorithm. Of course, a trade-off between cost and accuracy will have to be found.
T5.4 - Exascale optimization for AutoML: we develop optimization approaches in the automatic design of deep neural networks (DNNs) and the optimization of their hyper-parameters (AutoML), an important challenge. AutoML problems are more and more complex (e.g. dataset and network size) and their resource requirements in terms of computation and memory are ever increasing. Exascale optimization will allow to improve the accuracy of DNNs, reduce the energy and inference time, improve the robustness, and solve large-scale and/or complex learning tasks.

Deliverables	Title	Tasks	Dates
D5.1-RM-TR-P	Activity report included in annual report D0.2-TR	T5.*	M12,M24,M36,M48,M60
D5.2-S	Software package for optimization and shape optimization	T5.1, T5.2, T5.3	M36,M48,M60
D5.3-B	Benchmarking analysis report, including bottlenecks and breakthroughs	T5.*,T7.1	M12,M24,M36,M48,M60

Collaborations with NumPEx Projects: ExaSoft, Exa-DoST

WP Number		WP6	Lead	EP				Start	1	End	60
WP Title		Uncertainty Quantification									
Part. #	1	2	3	4	5						
Name	UNISTRA	CEA	INRIA	EP	SU						
PMs	10,2	66	0	147	0						

Objectives

Uncertainty Quantification (UQ) includes several steps, from uncertainty propagation to sensitivity analysis, to understand the important often correlated factors and then aim to reduce uncertainty through modeling improvements in WP1, finally establishing robust inversion or optimization under uncertainties. Enabling UQ in the different scales of a multiscale model remains a challenge that exascale computing will help address. The modeling of uncertainties may be strongly (e.g., in the case of stochastic equations) or weakly entangled into the models, calling for specific strategies but systematically requiring the evaluation of extremely high-dimensional integrals. Again, multi-fidelity stochastic or deterministic modeling will enable innovative tractable algorithms.

Description of work

T6.1 Kernel-based sensitivity analysis for high-dimensional data and integral computing
To handle very high-dimensional and multivariate data implied in exascale applications, tractable and relevant extensions of sensitivity analysis built upon kernel-based dependence measures will be developed. In addition, optimized computing schemes of high dimensional integrals will be investigated, in support of the uncertainty propagation step.

T6.2 UQ in a PDE solving framework The propagation of uncertainties (on the initial conditions, on the coefficients) in complex PDE solutions requires many simulations. Machine learning and stochastic spectral methods could provide suitable approximations of the solutions but their calibration can exceed by a few orders of magnitude the size of the underlying problems, making HPC strategies mandatory.

T6.3 Surrogate modeling for UQ Complex multi-physics and/or multi-scale problems often involve coupled, nested and chained numerical codes. The building and calibration of a global metamodel assembling all prediction uncertainties is a formidable task that requires HPC.

T6.4 Acceleration of the bricks of the UQ process steps by leveraging exascale calculations
The methodological developments of Tasks 6.1 to 6.3 will be integrated, in the opensource platforms Uranie and OpenTURNS dedicated to uncertainty quantification, taking advantage of the exascale computational properties. Benchmarking on exascale applications will be conducted.

Deliverables	Title	Tasks	Dates
D6.1.RM-TR-P	Activity report included in annual report D0.2-TR	T6.*	M24,M36,M48,M60
D6.2-S	Software release of URANIE T6.1 to T6.3	T6.*	M48,M60
D6.3.RB	Benchmarking analysis report, including bottlenecks and breakthroughs	T6.4;T7.1	M12,M24,M36,M48,M60

Collaborations with other NumPEx Projects:

WP Number		WP7	Lead	CEA			Start	1	End	
WP Title		Showroom, Benchmarking and Co-Design coordination								
Part. #	1	2	3	4	5					
Name	UNISTRA	CEA	INRIA	EP	SU					
PMs	112,2	0	144	60	48					

Objectives

WP7 objectives are *(i)* software development, from simple to advanced software testing including benchmarking to verify exascale capabilities and handling of identified challenges ***(B1-B13)*** and delivery of software packages in the framework proposed by ExaDIP in terms of CI/CD; *(ii)* the coordination of co-design activities within Exa-MA with ExaDIP; *(iii)* enabling a showroom of Exa-MA results and *(iv)* building training material from the results of Exa-MA.

WP7 builds on top of principles of non-regression, verification and validation. Within the framework of Exa-MA and more broadly of NumPEx, the various studies and developments carried out in the different Work Packages will have to be subjected to non-regression, verification and validation tests to be ensured before their integration into a demonstrator.

WP7 hosts the pool of engineers that will be recruited for these objectives. They will work at the crossroads of Exa-MA and the other projects, in particular ExaDIP. The management of WP7 will work in an Agile setting, following the project management plan set in WP0.

Description of work

T7.1 Testing and Benchmarking environment

Each relevant demonstrator must be identified in order to define a validation laboratory for the studies carried out in WPs 1 to 6. There will be different types of demonstrators: Level 1 demonstrator covers one to two WPs (e.g., AMR techniques); Level 2 demonstrator covers three to four WPs; Level 3 Demonstrator potentially covers all WPs. Some demonstrators, after benchmarking, will be retained by PC5 and will benefit from the processes set up in PC5 for the development and support of "mini apps".

T7.2 Co-design activities coordination

The integration process of the new developments will be based on all the demonstrators and will be broken down into non-regression, verification and validation processes. As a first approach, each demonstrator will have to deliver its current test process in order to guarantee its integrity: the new developments will not have to change the results acquired by these demonstrators except to change the options of the test case to benefit from the new developments. This process will guarantee the non-regression of the demonstrator while allowing to evaluate the new features and to compare them to the old ones.

The cases thus constructed (integrating via the input options the new developments) will be recorded as new non-regression, verification or validation tests depending on the level of the original tests. To these tests should be added new non-regression tests based solely on the functionality tested and not integrating the coupled options at this level. Verification and validation tests will be added on the same principle if it proves necessary to enrich the original base to better evaluate the contribution of the new methods.

T7.3 Showroom coordination

The presentation of the obtained results will be described in a unique format and presented in a dedicated web page. The results will be compared with the initial objectives in terms of performance and illustrated in the form of a figure of merit on the one hand, but also in the form of raw data (clock time, resources used, computer) and systematically compared with the initial performance of the demonstrator.

T7.4 Training

Training material on exascale toolboxes and min-apps are produced within the task.

Deliverables	Title	Tasks	Dates
D1.1-BR	Benchmarking analysis report, including bottlenecks and breakthroughs	T7.1	M12,M24,M36,M48,M60
D7.2-MR	co-design report	T7.2	M48,M60
D1.3-TR-P	Training material	T7.4	M12,M24,M36,M48,M60
D1.4-TR	Activity report included in annual report D0.2-TR	T7.*	M12,M24,M36,M48,M60

Collaborations with Projects: ExaSoft, ExaDoST, ExaAToW, ExaDiP

3. Project organisation and management

3.1. Projects managers

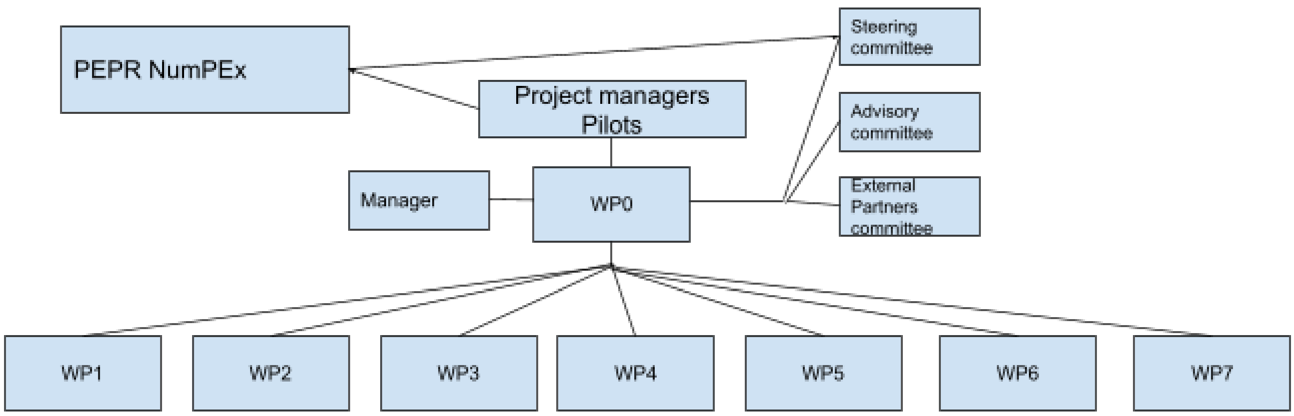
Exa-MA is managed by Christophe Prud'homme and Hélène Barucq, serving as strategic and operational managers in an interchangeable way to *(i)* represent the project in the PEPR and beyond and *(ii)* manage the project from the administrative, scientific and technical point of view. Both aspects are accompanied and advised by the steering committee of Exa-MA.

Christophe Prud'homme is a Professor of Applied Mathematics at IRMA, UMR 7501, at the University of Strasbourg. He has a rich experience in high-performance computing and has developed mathematical frameworks for efficient computation and PDE solutions to health applications. From 2000 to 2003, he worked at MIT, co-developing a mathematical framework for the reduced basis method, which led to much faster computations. From 2003 to 2006, he worked at EPFL, focusing on the development of mathematical and software frameworks for PDE solutions to health applications. He then became a Professor at the Université Joseph Fourier (Grenoble) and later at the University of Strasbourg. He has created a platform for collaborations between mathematicians and enterprises, MaiMoSiNE in Grenoble and Cemosis in Strasbourg. He is currently the French representative of the Center of Excellence Hidalgo2 funded by the PEPR and EU and member of the board of EU-Maths-IN the European association of national network for math-industry collaboration. His research at Cemosis is largely available in the C++ software library Feel++, which is widely used in industrial and research contexts as well as the simulation tool chain for the CNRS very large research infrastructure LNCMI for high field magnets.

Hélène Barucq is a Director of Research at INRIA. She is specialized in the construction of mathematical models and the development of advanced numerical methods for wave propagation. She is particularly interested in applications related to subsurface energy resources and for this, she works on the development of quantitative inversion methods adapted to anisotropic media. She is also working on the development of a computational environment dedicated to the study of stars, with an important focus on the Sun. High performance computing is a very important aspect of her work as it is essential to solve the large scale problems she considers. Her research activities are well integrated into the industrial environment, to the point that she leads an industrial research team bringing together academic researchers and researchers from TotalEnergies. She is also recognized in the academic world and regularly participates in national and European evaluation committees.

3.2. Organization of the partnership

The consortium is built in a balanced way, combining both theoretical and numerical expertise in mathematics and computer science in the different identified fields of the project.
The partners of Exa-MA have expertise in some or all WPs either as research contributors or as users. We join forces with our diverse and rich experiences, expertise and research contexts to tackle Exa-MA's challenges by setting collaborative actions in each WP (1-6) between at least two partners. In WP7, all partners contribute thanks to the pool of engineers offering a one-stop development, verification and deployment shop to Exa-MA. The richness of our partnership will foster new and effective software contributions.
Other than the project managers(or pilots), the key positions of the projects are the work package leaders for which WP1 WP2 and WP3 have three co-leaders while WP4, WP5, and WP6 have two. They are members of the steering committee of Exa-MA.
The Exa-MA governance is described by the figure below. The pilots are members of the PEPR steering committee. Exa-MA has *(i)* a steering committee composed of the pilots, representatives of the PEPR and the WP leaders *(ii)* a scientific advisory committee composed of international experts in the fields of Exa-MA, and *(iii)* the external partner board composed of industrial, public and private research infrastructures and academic colleagues to discuss scientific and technical collaborations and co-funding. The project management plan ***D0.1*** will detail the governance structures and responsibilities.



3.3. Management framework

The **management of the project** is taken care of within **Work Package 0 (WP0)**. We set an organizational structure that enables the execution of Exa-MA for : *(i)* the coordination and administration management and *(ii)* the scientific and technical management. All partners contribute through their respective representatives.

WP Number	**WP0**	**Lead**	**UNISTRA**	**Start**	**1**	**End**	**60**
WP Title	**Project Management**						

Part. #	1	2	3	4	5					
Name	UNISTRA	CEA	INRIA	EP	SU					
PMs	76,2	3,6	13,2	1,2	1,2					

Objectives

The work package has three objectives: Project Management, Technical and Scientific Management, and Administrative Management. Project Management focuses on communication, goal-oriented work, and daily management. Technical and Scientific Management oversees the technical direction of the project and integrates results. Administrative Management handles finances. The Consortium Agreement defines action governance and management procedures. WP0 will establish governance bodies and a Project Management Handbook that defines responsibilities and tools. The CA settles intellectual property with a focus on open-source software and a publication retention time policy if possible patents are identified.

Description of work

T0.1: Coordination and Administrative Management (Leader: UNISTRA) *Outputs: D0.1, D0.2*

This task is the overall coordination of the project. This includes communication management between the consortium and other projects, establishment and maintenance of project management bodies including the links with external partners, monitoring of risks, conflict management, identification and management of changes to the project plan, organization of project launch, and organization of audits. It also involves financial management of the project, monitoring financial forms such as cost statements, and organizing financial audits. This task aims to ensure the project's compliance with the Grant Agreement and Consortium Agreement and continuous monitoring of the work progress and resources.

T0.2: Scientific and Technical Management (Leader: UNISTRA) *Outputs: D0.2*

This task has three main fields of activity. It involves maintaining and monitoring the work plan, coordination and control of partners' results, and production of a roadmap document. The roadmap will define intermediate and final goals for the project's technological developments. The project management will be done through Github and an agile methodology will be used for engineering and development. The quality of deliverables will be ensured through an internal review process.

T0.3: Provision computing resources (Leader: INRIA) *Outputs: D0.2*

This task covers the provision of computing resources through close relationships with GENCI and the French supercomputing centers as well as the European ones. It requires coordination within NumPEx.

Task 0.4: Communication (Leader: INRIA) *Outputs: D0.2*

WP0 is responsible for synchronizing the communication between PC0 and the work packages of Exa-MA.

Deliverables	Title	Tasks	Dates
D0.1-TR	Project management Handbook	T0.1	M4
D0.2-TR	Project Management and Activity Report	T0.1,T0.2,T0.3,T0.4, T1*, T2*, T3*, T4* T5* T6* T7*	M12,M24,M36,M48,M60

The table below displays the Gantt diagram.

WP NUMBER	TASK TITLE	START DATE	DUE DATE	DURATION
Start Date	1/1/23			
0	**Mangement**			
T0.1	Coordination and Administrative Management	4/1/23	1/1/28	57
T0.2	Scientific and Technical Management	4/1/2023	1/1/2028	57
T0.3	Provision computing ressources	4/1/2023	3/1/2027	47
T0.4	Communication	4/1/2023	1/1/2028	57
1	**Discretization**			
T1.1	Geometry and Mesh at Exascale	10/1/23	10/1/25	24
T1.2	Adaptive Mesh Refinement for unstructured grids	10/1/23	1/1/28	51
T1.3	Adaptive Mesh Refinement for cartesian or block grids	10/1/24	1/1/28	39
T1.4	Finite Element Exascale Framework (FEEF)	10/1/23	1/1/28	51
T1.5	Exploit non conforming methods for efficient parallel computing	10/1/23	1/1/28	51
T1.6	Error control of time discretization for accurate space refinement	10/1/23	1/1/28	51
T1.7	Efficient multimodel/multphysics coupling	10/1/23	1/1/28	51
2	**Surrogate AI and ROM**			
T2.1	Surrogate models based on physics-driven Deep Learning	1/1/24	1/1/28	48
T2.2	PDE operator learning with Neural Operators	1/1/24	1/1/28	48
T2.3	Data-driven model order reduction	1/1/24	1/1/28	48
T2.4	Non-intrusive reduced basis methods for parametric problems	1/1/24	1/1/28	48
T2.5	Mixing low- and high-fidelity models	1/1/24	1/1/28	48
T2.6	Real-time models with super resolution methods	1/1/24	1/1/28	48
3	**Solvers**			
T3.1	Acceleration techniques for subspace-based methods	1/1/24	1/1/28	48
T3.2	Exploiting data-sparsity and multiple precision	1/1/24	1/1/28	48
T3.2.1	Modular, composable mixed precision Krylov solvers.	1/1/24	1/1/28	48
T3.2.2	Decoupling the data representation from the arithmetic: the variable accuracy paradigm	1/1/24	1/1/28	48
T3.2.3	Precision autotuning tools	1/1/24	1/1/26	24
T3.2.4	Silence data corruption and soft errors in iterative solvers	1/1/24	1/1/28	48
T3.3	Adaptive solution strategies for exascale multiphysical and multiscale models	8/1/24	10/1/25	14
T3.3.1	Multiphysics problems as saddle point problems	8/1/24	8/1/25	12
T3.3.2	Exascale resolution for simulations in partitioned coupling	8/1/24	8/1/25	12
4	**Inverse Problems and data assimilation**			
T4.1	Deterministic methods	10/1/23	10/1/27	48
T4.2	Stochastic methods	1/1/24	1/1/27	36
T4.3	Observations	10/1/23	10/1/27	48
T4.4	Taking advantage of multi-fidelity modeling	2/1/26	1/1/28	23
T4.5	Challenges of multi-fidelity in inverse problems: criteria to update reduced models	2/1/27	1/1/28	11
5	**Optimization**			
T5.1	Exascale combinatorial and continuous optimization	10/1/2023	1/1/2028	51
T5.2	Exascale surrogate-based optimization	10/1/2023	1/1/2028	51
T5.3	Exascale shape optimization	1/1/2024	1/1/2028	48
T5.4	Exascale optimization for AutoML mixed optimization	1/1/2024	1/1/2028	48
6	**Uncertainty Quantification**			
T6.1	Kernel-based sensitivity analysis for high-dimensional data and integral computing	10/1/2023	1/1/2028	51
T6.1.1	Kernel-based sensitivity analysis for high-dimensional and multivariate data	10/1/2023	10/1/2027	48
T6.1.2	Optimization of high-dimensional integral computing	10/1/2023	10/1/2027	48
T6.2	UQ in a PDE solving framework	10/1/2023	1/1/2028	51
T6.2.1	Intrusive UQ for PDE solving in a exascale framework	10/1/2023	10/1/2027	48
T6.3	Surrogate modeling for UQ	10/1/2023	1/1/2028	51
T6.3.1	UQ for surrogate models under physical constraints	10/1/2023	10/1/2027	48
T6.3.2	Tractability of Bayesian approaches for Gaussian Process Regression with high-dimensional inputs	10/1/2023	10/1/2027	48
T6.3.3	Metamodels for nested, chained and coupled codes	10/1/2023	10/1/2027	48
T6.4	Acceleration of the bricks of the UQ process steps by leveraging exascale calculations	1/1/2025	1/1/2028	36
T6.4.1	Integration in URANIE Platform and benchmarking on exascale applications	1/1/2025	1/1/2028	36
		1/1/2023	1/1/2023	0
7	**Demonstrators**			
T7.1	Benchmarking	10/1/2023	1/1/2028	51
T7.2	Co-design activities coordination	4/1/2023	1/1/2027	45
T7.3	Showroom coordination	10/1/2023	1/1/2028	51
T7.4	Training	10/1/2023	1/1/2028	51

The next tables display the milestones, the KPI and the critical risks of the execution of Exa-MA and how we plan to mitigate them respectively.

Table : Milestones

Milestone	Date	Description
MS1	**M9**	Identification of relevant application motifs
MS2	**M36,M48,M60**	Software stacks releases

Milestone	Date	Description
MS3	**M36,M48,M60**	Benchmarking at exascale
MS4	**M24,M36,M48,M60**	Releases of level1 demonstrators
MS4	**M30, M48, M60**	Releases of mini-apps
MS5	**M48,M60**	Releases of proxy-apps

Table : Key Performance indicators

KPI	Target
Number of Level 1 demonstrator at scale	At least 3 demonstrators per WP (except WP7) should be at scale
Number of Level 2 demonstrator at scale	At least 5 level 2 demonstrators should be at scale
Number of production applications including Exa-MA tools.	At least 5 mini-apps should be obtained by the end of the project and 4 proxy apps. About 20 libraries and software will be impacted by Exa-MA at different level of the project
Scale at which each Exa-MA software stack is validated.	Exa-MA software stack will be validated at full Exascale level.
Number of participants in training, tutorials, etc. presenting Exa-MA methodologies and software stack.	More than 100 trainees will have received Exa-MA related training in line with KPI-5 of PC0.
Number of international scientific publications with peer review process.	More than 50 papers will be published by Exa-MA participants during the project.
Contribution to KPI-7, KPI-8, KPI-9 and KPI-10 of PC0, mostly through software production.	Given the heterogeneity of contributions and in order to ensure uniformity between PCs, this KPI will be evaluated by PC0 that will ensure conformity with general NumPEx goals.

Table : Critical risks for implementation

Description of risk	Likelihood	Severity	WP	Proposed Risk-Mitigation Measures
Project Management Risks				
Partner related risks – underperforming, leaving the project, key-personnel temporarily not available	*Low*	*Medium*	WP0	Flexible project management structure and CA allow a quick shift of resources to alternative project partners and quick inclusion of new partners into the Consortium if needed. Furthermore, partners are involved in related areas, with more than a single staff member, ensuring an immediate substitution.
Planning problems: resources underestimated,	*Medium*	*Medium*	WP0	Potential solutions are involvement of other partners with available resources, rearrangement of resources among partners, change of the project plan within the

project timing not appropriate, further experts needed, deliverables/milestones delayed				self-assessment activities and EC, and ensuring timely implementation of corrective actions.
Collaboration issues – Consortium cannot agree, WP interaction not satisfactory, coordination not efficient	*Medium*	*High*	WP0	Project management provides appropriate decision-making and conflict resolution procedures which should be applied. As last instance, managements of the affected organizations, including the coordinating organization, will be involved in resolution.
New research results cause priority shifts.	*Low*	*Medium*	WP0	it may become necessary to change the project's direction due to dynamically changing situations, with a shift in priorities resulting
Scientific and Technical Risks				
Access to innovative HPC architectures cannot be provided	*Low*	*High*	ALL	The co-design activity with its benchmarking and optimization requires novel architectures to test.
Solutions developed do not scale or yield much lower performance improvements than envisioned	*Low*	*High*	ALL	Advancements made on exascale tools on scalability and optimization (including co-design activities) will be closely monitored by an outlined benchmarking and profiling strategy. Thus, Exa-MA can react early on when envisioned breakthroughs in scalability do not appear as projected.
Availability of compute resources and data does not suffice changed requirements	*Low*	*Medium*	ALL	Unavailability of compute resources
Delay in the availability of the infrastructure to be used for workflow development	*Low*	*Medium*	WP4	Unavailability of compute resources and complex workflow deployments.
Exploitation and Impact Risks				
Outreach to new communities and generated impact is lower than expected	*Low*	*High*	WP0	The consortium is setting up at the start of the project a reasonable communication strategy and compiles a roadmap for the campaigning activities including holding an EU Clustering event organized by Exa-MA. These objectives are revised on a regular basis to be adapted and further improved if needed.

3.4. Institutional strategy

UNISTRA has developed high performance computing capabilities for its demanding research application domains – nuclear fusion, astrophysics, chemistry, physics, biology, health and more recently quantum computing– and it is funding an interdisciplinary thematic institute (ITI) IRMIA++, since 2021, led by the Institut de Recherche en Mathématiques Avancées (IRMA) with HPC as a transverse topic. IRMA also led the Labex IRMIA from 2011 to 2021. In particular, this Labex regroups mathematicians(involved in PC1 Exa-MA), computer scientists (involved in PC2 ExaSoft) and applications domain researchers. Exa-MA is a formidable opportunity to push forward the frontiers in science and computing for UNISTRA and it is in line with its strategic agenda.

CEA plans to focus on the development and implementation of high-end computing resources and simulation software for a variety of applications including the evolution of neutron populations, flows, heat transfers, and mechanics. The CEA Energy Division has a diverse range of interests and the current trend is towards coupling multiple physics at different scales. Exaflop computers represent a significant opportunity to produce reference solutions, validate experiments, and improve understanding of complex phenomena at the finest scales.

Inria has identified high-performance computing and exascale as a priority research area in its strategic plan, and its teams contribute to the entire spectrum of challenges in exascale computing through software development,scientific and methodological activities. They have been involved in national, European, and international exascale initiatives, as well as the GENCI and PRACE programs. Inria also contributes to the transfer of knowledge to industry and aims to use HPC technology in an interdisciplinary way to address societal challenges, and a few teams will be involved in Exa-MA, working on the design of mathematical models and algorithms for extreme scale computing.

Ecole polytechnique is committed to the development of High Performance Computing (HPC), both in research and teaching. Research teams are directly involved in different themes that can be found in the Exa-MA project, such as uncertainty quantification. The HPC@Maths initiative aims at developing a competence in HPC using mathematical algorithms for computing and HPC, through a virtuous circle of research, training, and partnerships.

Sorbonne Université has been involved in high-performance computing for many years, with two research labs, LJLL and LIP6, that specialize in applied mathematics and computer science, respectively. Researchers from these labs have made many contributions to numerical libraries used in the HPC community, such as FreeFEM, MUMPS, HPDDM or PETSc. The PEPR's roadmap aligns with the ongoing projects, research, and teaching being carried out at both LJLL and LIP6, including the EUMaster4HPC.

We now indicate the details of the multi-year commitment of the partner institutions and then the expected budget ventilated in each partner institution. Researchers from each partner have been using GENCI resources continuously for the past ten years and are still working closely with the engineering teams of some national centers such as IDRIS and TGCC.

UNISTRA funds through the ITI IRMIA++ and the Labex IRMIA since 2021, 2 to 3 PhD fellowships per year, 2 postdoctoral fellowships and 2 engineer positions and most of them include HPC and many topics of Exa-MA. In particular, IRMIA++ funds a 2-year engineer position in the context of NumPEx as a collaboration between Exa-MA and PC2 ExaSoft colleagues from Strasbourg.

CEA The Energy Division, through its Simulation program, finances annually and in a stable manner over the next 5 years the activity of 300 full-time employees, to which must be added approximately 70 doctoral contracts. In this context, the part specifically dedicated to work on (very) high performance computing represents the activity of about 30 full-time employees, for 3 to 5 theses per year on this particular theme. As far as material resources are concerned, the Energies Department prefers to use the centralized resources of GENCI and the CCRT, via

collective agreements for an annual investment of about 150 k€ or specific projects based on DARI calls.
Inria has 221 project teams, 10 of which are clearly focused on HPC. In addition, more than 70 teams work on the modelling and simulation of large-scale complex systems and contribute to open parallel software packages for scientific computing available to the scientific community. These software packages are often supported by Inria's SED (Service d'Expérimentation et de Développement) engineers who ensure that the software is state of the art in terms of implementation, management, validation and deployment. On average, about 70 Ph.D theses are defended each year fon or with HPC, often funded or co-financed by industrial partners and awarded with the C3I label
Ecole polytechnique finances the HPC@Maths initiative through engineer and teacher-researcher positions. It also contributes to the implementation of shared computing resources, of a Computing and Data Mesocenter and a JupyterHub.
Sorbonne Université has ongoing and lasting relationships with industrial partners such as EDF and IFPEN, which on average fund 2 PhD fellowships per year on topics related to HPC. It finances a computing platform (MeSU) delivering more than 35 million core-hours each year to various departments of SU. Technical support could be obtained through the Service Unit of Sorbonne Université, SUMMIT, which can easily provide engineers manpower to accelerate developments.

4. Expected outcomes of the project

The expected scientific results of the project correspond directly to the objectives ***(O1-O5): (R1)* methods, algorithms, and implementations that, taking advantage of the exascale architectures,** empower modeling, solvers, model with data assimilation, optimization and uncertainty quantification, at levels that are unreachable at present; ***(R2)* open-source scientific software libraries** allowing to assemble specific critical reusable components, hiding the hardware complexity and exposing only the specific methodological interface ***(R3}* Methodological and Algorithmic Patterns** at exascale that can be reused efficiently in large-scale applications; ***(R4)* enable AI algorithms** to attain performances at exascale, exploiting the methods (O1) and the libraries (O2) developed; and ***(R5)* demonstrators**.

The impacts of the results are that they will ***(I1)*** design or revisit existing mathematical methods possibly AI enabled; ***(I2)*** contribute to the software development kits available for exascale applications identified in PC5-ExaDip and tackle the global challenges ***(C1-C9)***; ***(I3)*** contribute to open-source software available to the HPC community in France, Europe and beyond; ***(I4)*** bringing together our community to contribute solutions to scientific, societal and environmental challenges; ***(I5)*** creating/strengthening collaborations between external partners and Exa-MA partners and ***(I6)*** better structuring our community to respond effectively to European and international calls for projects.
The results will be available as reviewed publications in journals, reports, conferences, open-source software and development kits, training for initial and continuous formations and industrial transfer through industrial external partners.

Annexes/Appendices

Advisory board members

We hereby list suggestions for the advisory board members

WP1

- Christophe Geuzaine (Université de Liège, Belgium)
- Jan Hesthaven (CH,EPFL)

As Alternates:

- Jean-Francois Remacle (Belgium)
- Alexander Duester (Germany)
- Steven Owen (Sandia, USA)
- Christian Allen (UK)
- Rolf Krause (USI)
- Scott Canann (Siemens, USA)
- Suzanne M. Shontz (US)
- Nikos Chrisochoides (Old Dominion, US)

WP2

- Bernard Haasdonk (DE, U Stuttgart)
- Jan Hesthaven (EPFL)
- Karen Wilcox (Oden institut, Texas Usa)
- Gianluigi Rozza (Sissa, italy)
- Siddhartha Mishra (ETH, Suisse)
- Michael W. Mahoney (Berkeley, Usa)

WP3

- Axel Klawonn (Cologne),
- Lois McInnes (ANL),
- Ulrike Meier-Yang (Livermore)
- Wim Vanroose (Antwerpen Univ) (https://www.uantwerpen.be/en/staff/wim-vanroose/
- Mike Heroux (USA, Sandia Nat Lab, Trilinos)
- Jed Brown (University of Colorado at Boulder, PETSc, jed.brown@colorado.edu)
- Ivo Kabadshow (FZ-Juelich, i.kabadshow@fz-juelich.de)

WP4

- Omar Ghattas (UT Austin)
- Lars Nerger (Alfred Wegener Institute)

WP5

- Pascal Bouvry, Univ. of Luxembourg, Luxembourg
- Albert Zomaya, Univ of Sydney, Australia
- Marco Verani (http://www1.mate.polimi.it/~verani/) - shape optimisation

As alternates:

- Grégoire Danoy, University of Luxembourg, Luxembourg
- Enrique Alba, University of Malaga, Spain
- Jean-Paul Watson, Sandia National Laboratories, USA
- Ole Sigmund (https://orbit.dtu.dk/en/persons/ole-sigmund) - shape optimisation

WP6

- Bruno Sudret, ETH
- Corentin Lapeyre, CERFACS

As alternates:

- Serge Guillas, UCL
- Régis Lebrun , Airbus Group
- Yann Richet, IRSN
- Vincent Heuveline, Heidelberg

Bibliography

[CGAL/MMG] https://www.cgal.org/ and https://www.mmgtools.org/

[Cockburn2016] Bernardo Cockburn, Daniele Di Pietro, Alexandre Ern. Bridging the Hybrid High-Order and HybridizableDiscontinuous Galerkin Methods. ESAIM: Mathematical Modelling and Numerical Analysis, 2016, Polyhedral discretization for PDE, 50 (3), pp.635-650

[Bernardi2005] BERNARDI, Christine, MADAY, Yvon, et RAPETTI, Francesca. Basics and some applications of the mortar element method. GAMM‐Mitteilungen, 2005, vol. 28, no 2, p. 97-123

[Maday2001] Yvon Maday, Gabriel Turinici. Résolution d'EDP par un schéma en temps < pararéel >. Comptes rendus de l'Académie des sciences. Série I, Mathématique, 2001, 332 (7), pp.661-668.

[Buffa2012] Annalisa Buffa, Yvon Maday, Anthony T. Patera, Christophe Prud'Homme, Gabriel Turinici. A priori convergence of the Greedy algorithm for the parametrized reduced basis method. ESAIM: Mathematical Modelling and Numerical Analysis, 2012, 46 (3), pp.595-603.

[Barrault2004] Barrault, Maxime ; Maday, Yvon ; Nguyen, Ngoc Cuong ; Patera, Anthony T.An `empirical interpolation' method: application to efficient reduced-basis discretization of partial differential equations. C. R. Math. Acad. Sci. Paris 339 (2004), no. 9, 667--672.

[Spillane2014] Spillane, N., Dolean, V., Hauret, P., Nataf, F., Pechstein, C., & Scheichl, R. (2014). Abstract robust coarse spaces for systems of PDEs via generalized eigenproblems in the overlaps. Numerische Mathematik, 126, 741-770.

[Jolivet2014] Jolivet, P., & Nataf, F. (2014). HPDDM: high-Performance Unified framework for Domain Decomposition methods, MPI-C++ library.

[IOOSS2015] IOOSS, Bertrand et LEMAÎTRE, Paul. A review on global sensitivity analysis methods. Uncertainty management in simulation-optimization of complex systems, 2015, p. 101-122.

[Koliesnikova2021] D. Koliesnikova, I. Ramière and F. Lebon. A unified framework for the computational comparison of adaptive mesh refinement strategies for all-quadrilateral and all-hexahedral meshes: Locally adaptive multi- grid methods versus h-adaptive methods, Journal of Computational Physics, volume 437, 110310, 2021, doi:10.1016/j.jcp.2021.110310.

[Koliesnikova2022] D. Koliesnikova, I. Ramière and F. Lebon. Fully automatic multigrid adaptive mesh refinement strategy with controlled accuracy for nonlinear quasi-static problems, Computer Methods in Applied Mechanics and Engineering, volume 400, 115505, 2022, doi:10.1016/j.cma.2022.115505

[Raissi2019] M. Raissi, P. Perdikaris and G. E. Karniadakis, “Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations”, Journal of Computational Physics, Vol. 378, pp. 686-707 (2019)

[Cuomo2022] S. Cuomo, V. Schiano Di Cola, F. Giampaolo, G. Rozza, M. Raissi and F. Piccialli, “Scientific machine learning through physics–informed neural networks: where we are and what’s next”, Journal of Scientific Computing, Vol. 88, Art. No. 22 (2022)

[Wang2021] S. Wang, Y. Teng, P. Perdikaris, “Understanding and mitigating gradient flow pathologies in physics-informed neural networks”, SIAM Journal on Scientific Computing, Vol. 43, NO. 5, pp. 3055–A3081 (2021)

[Hu2022] Z. Hu, A.D. Jagtap, G.E. Karniadakis and K. Kawaguchi, “When do extended physics-informed neural networks (XPINNs) improve generalization?”, SIAM Journal on Scientific Computing, Vol. 44, No. 5, pp. A3158–A3182 (2022)

[Mishra2022] S. Mishra, R. Molinaro, “Estimates on the generalization error of physics-informed neural networks for approximating a class of inverse problems for PDEs”, IMA Journal of Numerical Analysis, Vol. 42, No. 2, pp. 981–1022 (2022)

[PINNs2022] Must-read papers on physics-informed neural networks. Github repository: https://github.com/idrl-lab/PINNpapers#papers-on-pinn-models

[Kovachki2021] N. Kovachki, Z. Li, B. Liu, K. Azizzadenesheli, K. Bhattacharya, A. Stuart, A. Anandkumar, “Neural Operator: learning maps between function spaces”, arXiv:2108.08481 (2021) https://arxiv.org/abs/2108.08481

[Li2021] Z. Li, H. Zheng, N. Kovachki, D. Jin, H. Chen, B. Liu, K. Azizzadenesheli, A. Anandkumar, “Physics-Informed neural operator for learning partial differential equations”, arXiv:2111.03794 (2021) https://arxiv.org/abs/2111.03794

[Lu2021] L. Lu, P. Jin, Z. Zhang and G.E. Karniadakis, “Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators”, Nature Machine Intelligence, Vol. 3, pp. 218–229 (2021)

[Li_22021] Z. Li, N. Kovachki, K. Azizzadenesheli, B. Liu, K. Bhattacharya, A. Stuart, A. Anandkumar, “Fourier Neural Operator for parametric partial differential equations”, arXiv:2010.08895 (2021) https://arxiv.org/abs/2010.08895

[Lee2020] K. Lee, K. T. Calberg,”Model reduction of dynamical systems on nonlinear manifolds using deep convolutional autoencoders”, Journal of Computational Physics, Vol. 404, No. 108973 (2020)

[Asch2022] M. Asch, “A toolbox for digital twins: from model-based to data-driven”, SIAM (2022)

[Asch2016] M. Asch, M.Bocquet and M. Nodet, “Data assimilation: methods, algorithms and applications”, SIAM (2016)

[Peherstorfer2018] B. Peherstorfer, K. Willcox, M. Gunzburger, “Survey of multifidelity methods in uncertainty quantification, inference and optimization”, SIAM Review, Vol. 60, No. 3, pp. 550–591 (2018)

[Saharia2022] C. Saharia, J. Ho, W. Chan, T. Salimans, D. J. Fleet and M. Norouzi, "Image super-resolution via iterative refinement," IEEE Transactions on Pattern Analysis and Machine Intelligence, to appear (2022). doi: 10.1109/TPAMI.2022.3204461

[Hao2021] H. Gao, L. Sun and J. X. Wan, “Super-resolution and denoising of uid ow using physics-informed convolutional neural networks without high-resolution labels”, Physics of Fluids, Vol. 33, No. 073603 (2021)

[Chakir2019] R. Chakir, Y. Maday and P. Parnaudeau, “A non-intrusive reduced basis approach for parametrized heat transfer problems”, Journal of Computational Physics, Vol. 376, pp. 617-633 (2019)

[Grosjean2022] E. Grosjean and Y. Maday, “Error estimate of the non-Intrusive reduced basis (NIRB) two-grid method with parabolic equations”, arXiv:2211.08897 (2022) https://arxiv.org/abs/2211.08897

[Lions2001] J.L. Lions, Y. Mada and G. Turinici, “Résolution d'EDP par un schéma en temps «pararéel»”, Comptes Rendus de l'Académie des Sciences-Series I-Mathematics, Vol. 332, No. 7, pp. 661-668 (2001)

[Maday2020] Y. Maday, Olga Mula. An Adaptive Parareal Algorithm. Journal of Computational and Applied Mathematics, Elsevier, 2020

[Hild2020] R. Hild, Optimization and control of high fields magnets, PhD thesis, Université de Strasbourg, November 2020.

[Debray2019] F. Debray, C. Grand-Clément, O. Jay, R. Hild, C. Prud'Homme, and C. Trophime, “Integrating modeling toolchain into high field Magnet monitoring and control system”, International Conference on Magnet Technology MT26, Vancouver, Canada (2019).

[Agullo2022] Emmanuel Agullo, Olivier Coulaud, Luc Giraud, Martina Iannacito, Gilles Marait, et al.. The backward stable variants of GMRES in variable accuracy. [Research Report] RR-9483, Inria. 2022, pp.1-77. ⟨hal-03776837v3⟩

[Agullo2018] Emmanuel Agullo, Eric Darve, Luc Giraud, Yuval Harness, Low-Rank Factorizations in Data Sparse Hierarchical Algorithms for Preconditioning Symmetric Positive Definite Matrices, *SIAM Journal on Matrix Analysis and Applications*, Society for Industrial and Applied Mathematics, 2018, 39 (4), pp.1701-1725

[Agullo2019] Emmanuel Agullo, Luc Giraud, Louis Poirel. Robust preconditioners via generalized eigenproblems for hybrid sparse linear solvers. *SIAM Journal on Matrix Analysis and Applications*, Society for Industrial and Applied Mathematics, 2019, 40 (2), pp.417-439. ⟨10.1137/17M1153765⟩

[Agullo2020a] Emmanuel Agullo, Franck Cappello, Sheng Di, Luc Giraud, Xin Liang, Nick Schenkels, Exploring variable accuracy storage through lossy compression techniques in numerical linear algebra: a first application to flexible GMRES, [Research Report] RR-9342, Inria Bordeaux Sud-Ouest. 2020

[Agullo2020b] Emmanuel Agullo, Siegfried Cools, Emrullah Fatih-Yetkin, Luc Giraud, Nick Schenkels, Wim Vanroose, On soft errors in the conjugate gradient method: sensitivity and robust numerical detection *SIAM Journal on Scientific Computing*, Society for Industrial and Applied Mathematics, 2020, 42 (6), ⟨10.1137/18M122858X⟩

[AlDaas2021] Hussam Al Daas, Laura Grigori, Pierre Jolivet, Pierre-Henri Tournier, A Multilevel Schwarz Preconditioner Based on a Hierarchy of Robust Coarse Spaces, *SIAM Journal on Scientific Computing,* 2021. https://doi.org/10.1137/19M1266964

[Amestoy2022a] Patrick Amestoy, Alfredo Buttari, Nicholas J Higham, Jean-Yves L'Excellent, Théo Mary, Bastien Vieublé, Combining sparse approximate factorizations with mixed precision iterative refinement. *ACM Transactions on Mathematical Software* https://hal.archives-ouvertes.fr/hal-03536031

[Amestoy2022b] Patrick Amestoy, Olivier Boiteau, Alfredo Buttari, Matthieu Gerest, Fabienne Jézéquel, et al.. Mixed Precision Low Rank Approximations and their Application to Block Low Rank LU Factorization. *IMA Journal of Numerical Analysis*, 2022. https://hal.archives-ouvertes.fr/hal-03251738

[Barr2006] Kenneth C. Barr and Krste Asanović. 2006. Energy-aware lossless data compression. ACM Trans. Comput. Syst. 24, 3 (August 2006), 250–291. https://doi.org/10.1145/1151690.1151692

[Blanchard2020a] Pierre Blanchard, Nicholas J Higham, Théo Mary. A Class of Fast and Accurate Summation Algorithms. *SIAM Journal on Scientific Computing*, 2020. https://hal.archives-ouvertes.fr/hal-02491081

[Blanchard2020b] Pierre Blanchard, Nicholas J Higham, Florent Lopez, Théo Mary, Srikara Pranesh. Mixed Precision Block Fused Multiply-Add: Error Analysis and Application to GPU Tensor Cores. *SIAM Journal on Scientific Computing*, 2020. https://hal.archives-ouvertes.fr/hal-02491076

[Bourgat1988] Bourgat, J-F., Roland Glowinski, and Patrick Le Tallec. Variational formulation and algorithm for trace operator in domain decomposition calculations. INRIA, 1988.

[Carson2022] Erin Carson, Noaman Khan, Mixed Precision Iterative Refinement with Sparse Approximate Inverse Preconditioning, to appear in *SIAM Journal on Scientific Computing*, 2022. https://arxiv.org/abs/2202.10204

[Collange2015] Sylvain Collange, David Defour, Stef Graillat and Roman Iakymchuk. Numerical reproducibility for the parallel reduction on multi- and many-core architectures. *Parallel Comput.*, 49:83–97, 2015.

[Chen2003] CHEN, Mo; FOWLER, Mark L. The importance of data compression for energy efficiency in sensor networks. In: Conference on Information Sciences and Systems. 2003. p. 13.

[Connolly2022] M. Connolly, N. J. Higham, S. Pranesh, *Randomized Low Rank Matrix Approximation: Rounding Error Analysis and a Mixed Precision Algorithm*, 2022. http://eprints.maths.manchester.ac.uk/2863/

[Cools2018] Siegfried Cools, Emrullah Fatih Yetkin, Emmanuel Agullo, Luc Giraud, Wim Vanroose. Analyzing the Effect of Local Rounding Error Propagation on the Maximal Attainable Accuracy of the Pipelined Conjugate Gradient Method, *SIAM Journal on Matrix Analysis and Applications*, Society for Industrial and Applied Mathematics, 2018, 39 (1), pp.426 – 450. ⟨10.1137/17M1117872⟩

[Coulaud2022] A robust GMRES algorithm in Tensor Train format.Olivier Coulaud, Luc Giraud, Martina Iannacito, [Research Report] RR-9484, Inria. 2022, pp.1-48. ⟨hal-03776529v3⟩

[Farhat1991] Farhat, C. and Roux, F.-X. (1991), A method of finite element tearing and interconnecting and its parallel solution algorithm. Int. J. Numer. Meth. Engng., 32: 1205-1227. https://doi.org/10.1002/nme.1620320604

[Giraud2022] A block minimum residual norm subspace solver with partial convergence management for sequences of linear systems, Luc Giraud, Yan-Fei Jing, Yanfei Xiang, *SIAM Journal on Matrix Analysis and Applications*, Society for Industrial and Applied Mathematics, 2022, 43 (2), pp.710-739. ⟨10.1137/21m1401127⟩

[Giraud2007] Convergence in backward error of *relaxed* GMRES. L. *Giraud*, S. Gratton, and J. Langou. *SIAM* J. Sci. Comput., 29(2):710–728, 2007.

[Bourras2000] A relaxation strategy for inner- outer linear solvers in domain decomposition methods. Bouras, V. Frayss e, and L. Giraud.. Technical Report TR/PA/00/17, CERFACS, 2000.

[Graillat2019] S. Graillat, F. Jézéquel, R. Picot, F. Févotte, B. Lathuilière. Auto-tuning for floating-point precision with Discrete Stochastic Arithmetic, Journal of Computational Science, 36, 2019. https://hal.archives-ouvertes.fr/hal-01331917

[Graillat2022] Stef Graillat, Fabienne Jézéquel, Théo Mary, Roméo Molina. Adaptive precision sparse matrix-vector product and its application to Krylov solvers. 2022. https://hal.archives-ouvertes.fr/hal-03561193

[Higham2021] Nicholas Higham and Theo Mary. Solving Block Low-Rank Linear Systems by LU Factorization is Numerically Stable, IMA J. Numer. Anal., (2021).

[Iakymchuk2022] Roman Iakymchuk, Maria Barreda Vayá, Stef Graillat, José I Aliaga and Enrique S Quintana-Ortí. Reproducibility of parallel preconditioned conjugate gradient in hybrid programming environments. *The International Journal of High Performance Computing Applications*, 34(5):502–518, 2020.

[Jezequel2021] F. Jézéquel, S. sadat Hoseininasab, T. Hilaire, Numerical validation of half precision simulations, 9th World Conference on Information Systems and Technologies, 2021. https://hal.archives-ouvertes.fr/hal-03138494

[Lelewer1987] Debra A. Lelewer and Daniel S. Hirschberg. 1987. Data compression. ACM Comput. Surv. 19, 3 (Sept. 1987), 261–296. https://doi.org/10.1145/45072.45074

[Ma2022] Chupeng Ma, Robert Scheichl, and Tim Dodwell, Novel Design and Analysis of Generalized Finite Element Methods Based on Locally Optimal Spectral Approximations, SIAM Journal on Numerical Analysis 2022 60:1, 244-273

[Mukunoki2016] Daichi Mukunoki, Toshiyuki Imamura, Reduced-Precision Floating-Point Formats on GPUs for High Performance and Energy Efficient Computation, 2016 IEEE International Conference on Cluster Computing (CLUSTER).

[Nataf2021] Frédéric Nataf, Pierre-Henri H Tournier. A GenEO Domain Decomposition method for Saddle Point problems. 2021. ⟨hal-02343808v5⟩

[Paige2006] Christopher C. Paige, Miroslav Rozlozník, and Zdenvek Strakos, Modified Gram-Schmidt (MGS), Least Squares, and Backward Stability of MGS-GMRES, *SIAM Journal on Scientific Computing* 2006. https://doi.org/10.1137/050630416

[Sousedík2022] Sousedík, B. Inexact and primal multilevel FETI-DP methods: a multilevel extension and interplay with BDDC. Int J Numer Methods Eng. 2022; 123(20): 4844- 4858. doi:10.1002/nme.7057

[Viot2018] Solving coupled problems of lumped parameter models in a platform for severe accidents in nuclear reactors, L. Viot, L., Saas, & F. De Vuyst, International Journal for Multiscale Computational Engineering, 16(6), 555–577, 2018

[Feel++2012] Feel++: A computational framework for Galerkin methods and advanced numerical methods. In : ESAIM: Proceedings. EDP Sciences, 2012. p. 429-455.

[Jolivet2013] Pierre Jolivet, Frédéric Hecht, Frédéric Nataf, Christophe Prud'homme. Scalable domain decomposition preconditioners for heterogeneous elliptic problems. SC13. https://hal.archives-ouvertes.fr/hal-00939957/document

[Arridge2019] S. Arridge, P. Maass, O. Oktem, C. Schonlieb. *Solving inverse problems using data-driven models.* Acta Numerica (2019), pp. 1-174.

[Bartz2020] T. Bartz-Beielstein, B. Filipic, P. Korosec, E-G. Talbi (Eds.), "High-performance simulation based optimization", Springer, 2020.

[Talbi2006] E-G. Talbi (Ed.), "Parallel Combinatorial Optimization", Wiley, ISBN-978-0-471-72101-7, 2006.

[Brevault2020] L. Brevault, J. Pelamatti, A. Hebbal, M. Balesdent, E-G. Talbi, N. Melab. Multi-disciplinary design optimization related issues: multi-objective and mixed continuous/discrete optimization. Aerospace system analysis in uncertainty, edited by L. Brevault, M. Balesdent, J. Morio, Springer, pp.321-358, 2020.

[Pelamatti2020] J. Pelamatti, L. Brevault, M. Balesdent, E-G. Talbi, Y. Guerin. Overview and comparison of Gaussian Process-based surrogate models for mixed continuous and discrete variables: application on aerospace design problems. High Performance simulation based optimization, edited by T. Bartz-Beielstein, B. Filipic, P. Korosec, E-G. Talbi, Springer, pp.189-224, 2020.

[Talbi2021] E-G. Talbi, Automated design of deep neural networks: A survey and unified taxonomy, ACM Computing Surveys, Vol.54, No.2, 2021.

[Talbi2022] E-G. Talbi, «Machine learning into metaheuristics: A survey and taxonomy of data-driven metaheuristics», ACM Computing Surveys, Vol.54, No.4, 2021.

[Sadet2022] J. Sadet, F. Massa, T. Tison, E-G. Talbi, I. Turpin, «Deep Gaussian process for the approximation of a quadratic eigen value problem: application to friction-induced vibration», Vibration, Vol.5, No.2, 2022.

[Habbal2021] Hebbal, L. Brevault, M. Balesdent, E-G. Talbi, N. Melab «Multi-fidelity modelling with different input domain definitions using deep Gaussian processes», Structural & Multidisciplinary Optimization Journal, Vol.63, N.5, pp.2267-2288, 2021.

[Talbi2019] E-G. Talbi, “A unified view of parallel multi-objective evolutionary algorithms”, Journal of Parallel and Distributed Computing JPDC, vol.133, pp.249-358, 2019.

[Nakib2019] A.Nakib, L. Souquet, E-G. Talbi, “Parallel fractal decomposition based algorithm for big continuous optimization problems”, Journal of Parallel and Distributed Computing, Vol.133, pp.297-306, 2019.

WP6

[Da Veiga2015] Da Veiga, S. (2015). Global sensitivity analysis with dependence measures. Journal of Statistical Computation and Simulation.

[Da Veiga2021] Da Veiga, S., Gamboa, F., Iooss, B., and Prieur, C. (2021). Basics and Trends in Sensitivity Analysis: Theory and Practice in R. Society for Industrial and Applied Mathematics.

[El Amri2021] El Amri, M. R. and Marrel, A. (2021). More powerful HSIC-based independence tests, extension to space-filling designs and functional data. Preprint available at https://hal-cea.archives-ouvertes.fr/cea-03406956.

[El Amri2022] El Amri, M. R. and Marrel, A. (2022). Optimized HSIC-based tests for sensitivity analysis: Application to thermalhydraulic simulation of accidental scenario on nuclear reactor. Quality and Reliability Engineering International.

[Fekhari2022] Fekhari, E., Iooss, B., Chabridon, V., Muré J.. Kernel-based quadrature applied to offshore wind turbine damage estimation. MASCOT-NUM 2022, 2022, Clermont-Ferrand, France. https://hal.archives-ouvertes.fr/hal-03713013

[Girard 2007]. Girard, S. and Iovleff, S. (2007) Auto-associative models, nonlinear Principal component analysis, manifolds and projection pursuit. LNCSE, Springer-Verlag.

[Uriarte2022] Carlos Uriarte, David Pardo, Ángel Javier Omella, A Finite Element based Deep Learning solver for parametric PDEs, Computer Methods in Applied Mechanics and Engineering, Volume 391, 2022

[Tenenbaum2000] Tenenbaum, Joshua B., Vin de Silva, and John C. Langford. A global geometric framework for nonlinear dimensionality reduction. science 290.5500 (2000): 2319-2323.

[Sapienza2018] Sapienza, Facundo & Jonckheere, Matthieu & Groisman, Pablo. (2018). Weighted Geodesic Distance Following Fermat's Principle.

[Pele2021] Kathleen Pele. Prédiction statistique de chemins de fissuration : application au matériau cimentaire. Mathématiques générales [math.GM]. Ecole Centrale Marseille, 2021

[Kriege2020] Kriege, N.M., Johansson, F.D. & Morris, C. A survey on graph kernels. Appl Netw Sci 5, 6 (2020)

[LeMaitre2022] O.P. Le Maître, el al., A Stochastic Galerkin method for uncertain time-fractional diffusion equations, (in preparation).

[Boukaram2020] W. Boukaram, et al., Hierarchical Matrix Approximation for Space-Fractional Diffusion Equations, Computer Methods in Applied Mechanics and Engineering, 369, pp.113191, (2020).

[Venkovic2021] N. Venkovic, et al., Recycling Krylov subspace strategies for sequences of sampled stochastic elliptic equations, [Research Report] RR-9425, Inria Bordeaux – Sud Ouest. 2021

[Reis2021] J.F. Reis, O.P. Le Maître, P.M. Congedo and P. Mycek. Acceleration of Domain Decomposition Methods for Stochastic Elliptic Equations, Computer Methods in Applied Mechanics and Engineering, 381, pp.113845, (2021).

[Rasmussen2006] C. E. Rasmussen & C. K. I. Williams, Gaussian Processes for Machine Learning, the MIT Press, 2006

[Swiler2020] L. Swiler, M. Gulian, A. Frankel, C. Safta, and J. Jakeman, A Survey of Constrained Gaussian Process Regression: Approaches and Implementation Challenges, Journal of Machine Learning for Modeling and Computing, Volume 1, 2020, pp. 119-156

[Gu0218] M. Gu, X. Wang, and J. O. Berger, Robust Gaussian Stochastic Process Emulation, Ann. Statist., Volume 46, 2018, pp. 3038-3066

[Marque2020] S. Marque-Pucheu, G. Perrin, and J. Garnier, An efficient dimension reduction technique for the Gaussian process emulation of two nested codes with functional outputs, Computational Statistics, Volume 35, 2020, pp. 1059-1099