

Haut Conseil de l'Évaluation de la Recherche et
de l'Enseignement Supérieur



DOCUMENT D'AUTOÉVALUATION
Équipe PEQUAN



Campagne d'évaluation 2023-2024 — Vague D

Table des matières

1	INFORMATIONS GÉNÉRALES SUR L'ÉQUIPE PEQUAN	3
1.1	Les thématiques scientifiques et leurs enjeux	3
	Thématiques	3
	Positionnement scientifique par rapport au contexte international	3
	Avancées scientifiques majeures dans la période	5
	Animation scientifique	8
2	INTRODUCTION DU PORTFOLIO	9
3	AUTOÉVALUATION DU BILAN	10
3.1	Autoévaluation de l'équipe	10
	Domaine 2. Attractivité	10
	Domaine 3. Production scientifique	12
	Domaine 4. Inscription des activités de recherche dans la société	14
4	RÉFÉRENCES BIBLIOGRAPHIQUES EXTERNES	16
5	RÉFÉRENCES BIBLIOGRAPHIQUES SIGNIFICATIVES DE PEQUAN	17
A	ANNEXE — MEMBRES PERMANENTS AU 31/12/2022	20

1 INFORMATIONS GÉNÉRALES SUR L'ÉQUIPE PEQUAN

Nom de l'équipe : PErformance et QUalité des Algorithmes Numériques (PEQUAN)

Responsable de l'équipe : Fabienne Jézéquel

	2017	2018	2019	2020	2021	2022
PR	3	3	3	2	2	2
MCF HDR	2	2	3	3	2	2
MCF	4	4	2	2	2	2
DR	0	0	0	0	0	0
CR HDR	1	1	0	0	0	0
CR	1	1	1	2	1	1
Total permanents	11	11	9	9	7	7
Émérites	1	1	1	1	1	0
Doctorants	8	5	1	1	2	6
Ingénieurs CDD ou hors tutelles	1	1	0	1	1	0
Post-doc, ATER, etc.	0	0	0	1	1	1
Stagiaires	4	4	2	2	4	4
Total non permanents	13	10	3	5	8	11
Total avec émérites	25	22	13	15	16	18
Equivalent temps plein recherche	6.5	6.5	5.0	5.5	4.0	4.0

TABLE 1 – Personnels PEQUAN sur la période 2017-2022 (au 1er juillet de chaque année)

Entre le 1er janvier 2017 et le 31 décembre 2022, l'équipe a accueilli 2 CR CNRS et perdu 7 membres permanents, la majorité des départs étant due à des mobilités géographiques.

1.1 Les thématiques scientifiques et leurs enjeux

Thématiques

Les activités de recherche de l'équipe PEQUAN visent à concevoir des algorithmes numériques à la fois performants et fiables. Ces recherches sont menées à bien grâce à la réunion de compétences en arithmétique des ordinateurs et en calcul numérique haute performance qui constitue une des spécificités de l'équipe. Les activités de PEQUAN portent notamment sur la résolution de problèmes dans une grande variété de domaines d'application. Nos principaux axes de recherche sont les suivants.

- ▶ fiabilité des algorithmes numériques : conception d'algorithmes efficaces et précis en arithmétique à virgule flottante ou à virgule fixe, ainsi que leur validation numérique
- ▶ calcul numérique haute performance : conception d'algorithmes parallèles et optimisation de leurs performances sans dégrader la précision de leurs résultats
- ▶ problèmes liés à l'assimilation des images : prise en compte des modèles physiques dans la résolution de problèmes inverses en imagerie, complétion des modèles par des approches hybrides d'apprentissage statistique et de d'assimilation variationnelle de données.

L'équipe PEQUAN participe aux activités de 2 axes du LIP6 : "Sécurité, sûreté et fiabilité" (SSR) et "Théorie et outils mathématiques pour l'informatique" (TMC).

Positionnement scientifique par rapport au contexte international

Analyse des erreurs commises. Les simulations numériques en virgule flottante peuvent générer à chaque opération arithmétique une erreur d'arrondi. L'accumulation de ces erreurs d'arrondi peut invalider les résultats calculés. Dans le domaine de la fiabilité numérique, l'équipe contribue à différentes approches, notamment l'analyse a priori de l'erreur commise, l'arithmétique d'intervalles et l'arithmétique stochastique. L'équipe a particulièrement développé son expertise sur cette dernière approche, probabiliste, depuis les travaux pionniers de Jean Vignes. À l'international, les travaux sur l'arithmétique d'intervalles sont notamment menés en Allemagne, aux États-Unis et au Japon. L'arithmétique d'intervalles consiste à effectuer les calculs avec des intervalles qui contiennent toujours le résultat exact. Les bornes fournies sont donc garanties, tandis que l'arithmétique stochastique produit une estimation de l'erreur d'arrondi commise. En revanche, l'arithmétique d'intervalles nécessite des algorithmes spécifiques, alors que l'arithmétique stochastique permet de contrôler l'erreur générée par tout type d'algorithme numérique.



Amélioration de la précision. En raison de la propagation des erreurs d'arrondi, certains problèmes doivent être résolus avec un format à virgule flottante plus large. C'est notamment le cas des applications qui effectuent des tâches complexes dans des domaines tels que la théorie des champs quantiques, la simulation de supernova, la physique des semi-conducteurs ou les calculs d'orbite planétaire. Étant donné que le calcul à l'échelle de l'Exascale (10^{18} opérations en double précision par seconde) a été atteint en 2020, l'obtention de résultats précis en arithmétique à virgule flottante sur de tels ordinateurs est un défi majeur.

L'équipe travaille sur l'amélioration de la précision des résultats. À cet effet, les calculs peuvent être menés dans une précision élevée, par exemple quadruple (128 bits) ou arbitraire via l'utilisation d'une bibliothèque telle MPFR¹. Des expansions flottantes ont aussi été proposées. Celles-ci consistent à représenter un nombre par une somme non évaluée de deux flottants. Les bibliothèques DDFUN et DQFUN développées à Berkeley² permettent des calculs respectivement avec l'association des flottants double-double ou double-quadruple. Par ailleurs, la bibliothèque MPLAPACK³ développée à RIKEN au Japon implante des algorithmes d'algèbre linéaire en précision élevée grâce à la bibliothèque MPFR ou aux expansions flottantes. Enfin, les méthodes compensées fondées sur les transformations sans erreur ou EFT permettent d'améliorer la qualité numérique d'algorithmes de base de l'algèbre linéaire sans changer la précision de calcul, au prix d'opérations arithmétiques supplémentaires. Ces dernières années, l'équipe a contribué au contrôle d'erreur en précision élevée ou arbitraire et au développement des méthodes compensées.

Reproductibilité numérique. Comme l'addition en virgule flottante en précision finie n'est pas associative, le résultat d'une sommation peut varier d'une machine parallèle à l'autre, voire d'une exécution à l'autre. Ces écarts s'aggravent sur des architectures hétérogènes, telles que des clusters composés de CPU standard en conjonction avec des GPU. Le non-déterminisme des calculs à virgule flottante dans les programmes parallèles pose des problèmes de validation et de débogage, et peut même conduire à des blocages. Ces problèmes deviendront de plus en plus critiques au fur et à mesure que la tendance vers des plateformes hétérogènes à grande échelle se poursuivra.

Des recherches sont donc menées pour proposer des bibliothèques de calcul scientifique permettant d'obtenir des résultats reproductibles. Ces bibliothèques reposant souvent sur du calcul d'algèbre linéaire, les recherches se sont surtout intéressées à rendre les BLAS (Basic Linear Algebra Subprograms) reproductibles. La première bibliothèque disponible a été ReproBLAS⁴ développée à Berkeley. On peut aussi citer la bibliothèque OzBLAS développée au Japon. L'équipe PEQUAN a proposé la bibliothèque ExBLAS. Il s'agit d'une bibliothèque basée sur l'usage d'un accumulateur long qui permet de calculer des sommes et des produits scalaires avec un arrondi au plus près. Cela permet de disposer de noyaux de calcul reproductibles, précis et efficaces.

Fiabilité des algorithmes embarqués. La thématique numérique s'applique aussi aux algorithmes embarqués, qui sont implémentés sur des cibles variées, que ce soient des processeurs "classiques" (plutôt ARM ou même RISC-V), des micro-contrôleurs ou des FPGA. Les problématiques rencontrées viennent principalement des contraintes de ressources et d'énergie. Les arithmétiques considérées peuvent être à virgule flottante ou à virgule fixe et, dans le cas du FPGA, les largeurs considérées peuvent être quelconques. Le problème posé est souvent de trouver une implémentation minimisant un "coût" de calcul (une surface de circuit, une consommation, une empreinte mémoire, une utilisation des ressources) sous contrainte de précision numérique. Dans le cas de l'implémentation matérielle (FPGA), le problème devient un problème d'allocation des largeurs (largeur en nombre de bits des différents opérateurs utilisés). Ce problème est souvent étudié spécifiquement pour des classes précises d'algorithmes (par exemple du traitement du signal ou de vidéo), mais aussi de manière plus générale. On peut notamment citer les travaux de l'Imperial College (Department of Electrical & Electronic Engineering) qui utilisent une analyse statique pour l'évaluation de la précision nécessaire pour la résolution du problème d'implémentation optimale.

Calculs en précision faible ou mixte. Les bibliothèques de calcul scientifique doivent intégrer des algorithmes efficaces, adaptés aux différentes architectures parallèles, voire repensés pour les machines Exascale. Aux États-Unis, le projet "Exascale Computing Project" (ECP, 2016-23) inclut un volet dédié au développement de bibliothèques scientifiques nommé xSDK (Extreme-Scale Scientific Software Development Kit). Celui-ci regroupe différents industriels et académiques dont Jack Dongarra (Prix Turing 2021). PEQUAN participe au PEPR NUMPEX (NUMérique Pour l'EXascale, programme d'investissements d'avenir, 2023-2028). Sa contribution porte sur la conception de méthodes et algorithmes pour l'Exascale. Les bibliothèques d'algèbre linéaire, fondées sur les BLAS, évoluent pour tenir compte des moyens de calcul actuels et à venir. Elles doivent répondre à différents défis : réduction du temps de calcul et des communications, passage à l'échelle dans les environnements multi-

1. <http://mpfr.org>

2. <https://www.davidhbailey.com/dhbsoftware>

3. <https://github.com/nakatamaho/mplapack>

4. <https://bebop.cs.berkeley.edu/reproblas/>



cœur, prise en compte des architectures hybrides (par exemple CPU-GPU). Certaines de ces bibliothèques bénéficient de contributions internationales, notamment Ginkgo, hypre, MAGMA, MUMPS, PETSc, PLASMA, SuperLU. L'équipe PEQUAN contribue au développement de deux de ces bibliothèques : MUMPS et PETSc.

L'utilisation d'une précision réduite ou mixte dans les programmes permet de gagner en performance grâce à une réduction du temps d'exécution, de la consommation énergétique et de l'empreinte mémoire. Le projet européen "Open Transprecision Computing" (OPRECOMP, 2017-21) a fédéré des recherches sur la précision mixte couvrant différents aspects matériels et logiciels. Aux États-Unis, l'action "xSDK multiprecision" au sein du projet ECP porte sur l'introduction de la précision mixte dans les bibliothèques scientifiques. L'équipe PEQUAN travaille sur la conception d'algorithmes et la génération de code en précision mixte. Ces travaux pourront se poursuivre au sein du PEPR NUMPEX.

Ces dernières années de nombreux travaux ont porté sur la conception d'algorithmes numériques en précision mixte, avec notamment des avancées majeures en algèbre linéaire, qui sont présentées dans le *survey* [Higham and Mary, 2022]. L'équipe PEQUAN a contribué à plusieurs de ces avancées, en particulier pour la résolution de systèmes linéaires par raffinement itératif [Amestoy et al., 2022c] et pour les approximations de rang faible [Amestoy et al., 2022a].

Par ailleurs, des outils d'ajustement automatique de la précision ont été développés. Ceux-ci peuvent proposer une configuration de types en précision mixte dans un programme sans connaissance préalable des algorithmes qui sont implantés. Différents outils tels FPTuner [3], Rosa/Daisy [4, 5] ou TAFFO [2] reposent sur une approche statique et ne sont pas destinés à être utilisés sur de très gros codes. D'autre part, des outils dynamiques tels CRAFT HPC [10], Precimonious [12], HiFPTuner [8], ADAPT [11] ou FloatSmith [9] ont été proposés pour l'auto-tuning de précision dans de grands codes HPC. L'équipe développe PROMISE⁵ [Graillat et al., 2019, Jézéquel et al., 2021], un logiciel dynamique d'auto-tuning de précision qui, grâce à un contrôle des erreurs d'arrondi, fournit une configuration de types validée.

Assimilation des images et apprentissage profond. Notre équipe s'était initialement positionnée sur l'utilisation de l'assimilation de données pour résoudre des problèmes inverses en traitement d'images principalement géophysiques (estimation de mouvement, inpainting, super-résolution). Avec le développement récent de l'apprentissage statistique profond, nous avons été la première équipe en 2020, [Filoche et al., 2020], à proposer des approches hybrides combinant assimilation variationnelle de données pour résoudre le problème inverse et l'apprentissage statistique pour estimer le modèle. A notre connaissance, très peu d'équipes travaillent sur ce sujet, principalement en France, [1, 6, 7]. Par ailleurs, les codes associés aux publications sont mis à disposition de la communauté.

Avancées scientifiques majeures dans la période

Analyse d'erreur par l'arrondi stochastique. Le développement des ressources de calcul va de pair avec une augmentation du nombre d'opérations arithmétiques, chacune pouvant entraîner une erreur d'arrondi. Il est donc crucial de valider numériquement les résultats produits et de s'assurer que l'erreur reste acceptable pour l'application visée. Nous avons montré à travers différents travaux récents comment répondre à ces défis par l'utilisation d'arrondi stochastique (*stochastic rounding* ou SR).

D'une part, nous avons montré que remplacer les modes d'arrondis classiques tel que l'arrondi au plus près par l'arrondi stochastique peut grandement améliorer la qualité des résultats. En effet, il a été conjecturé par Wilkinson en 1960 que les effets probabilistes lors de calculs en précision finie peuvent réduire la taille des bornes d'erreur d'un facteur \sqrt{n} pour la plupart des calculs d'algèbre linéaire (produit matriciel, résolution de système linéaire) impliquant des matrices d'ordre n . Nous avons récemment prouvé cette conjecture et identifié sous quelles conditions elle est applicable [Higham and Mary, 2020]. Par la suite, nous avons remarqué que ces conditions sont justement satisfaites par l'utilisation de l'arrondi stochastique [Connolly et al., 2021]. Ceci explique donc les bénéfices observés de ce mode d'arrondi, qui connaît un fort regain d'intérêt notamment en IA.

D'autre part, l'équipe développe l'Arithmétique Stochastique Discrète (ASD) qui permet d'estimer les erreurs d'arrondis dans les programmes. L'ASD consiste à effectuer chaque opération arithmétique plusieurs fois avec le mode d'arrondi stochastique. Un test statistique permet d'estimer le nombre de chiffres corrects des résultats. L'ASD est implantée dans la bibliothèque CADNA⁶ pour les calculs en précision classique (par exemple simple et/ou double) et dans la bibliothèque SAM⁷ pour les calculs en précision arbitraire. Ces dernières années, la bibliothèque CADNA a été étendue pour permettre le contrôle d'erreur dans les codes sur CPU en précision quadruple [Graillat

5. <http://promise.lip6.fr>

6. <http://cadna.lip6.fr>

7. <https://www-pequan.lip6.fr/~jezequel/SAM>



et al., 2018b] ou en précision demie (native ou émulée) [Jézéquel et al., 2021]. Les performances de la version de CADNA pour architectures hybrides CPU-GPU ont été améliorées et ses fonctionnalités étendues [Eberhart et al., 2018]. De plus, une version alternative de SAM est maintenant proposée pour les calculs en précision mixte. Afin de réduire le coût de la validation numérique, nous avons montré que, sous certaines conditions, plusieurs appels à des noyaux de calcul classiques peuvent se substituer à des calculs en ASD. Nos expérimentations numériques ont alors mis en évidence un gain en performance allant jusqu'à 3 ordres de grandeur [Jézéquel et al., 2020]. Des versions en ASD de l'algorithme de PGCD de polynômes et de la méthode de Newton qui tiennent compte du bruit numérique ont été proposées pour le calcul de racines multiples de polynômes, particulièrement instable [Graillat et al., 2021]. Ces nouveaux algorithmes ont permis d'obtenir des résultats plus fiables et plus performants (de plusieurs ordres de grandeur en temps d'exécution) par rapport aux algorithmes produits par l'équipe en 2016 pour le même type de problème.

Amélioration de la précision et de la reproductibilité des résultats. L'équipe travaille aussi sur l'amélioration de la qualité numérique ou de la reproductibilité des résultats. Les méthodes compensées sont fondées sur des transformations dites "sans erreur" (EFT) et permettent d'améliorer la précision des résultats en tenant compte des erreurs d'arrondi commises. Elles ont été conçues pour l'arrondi au plus près et ne sont plus exactes en arrondi dirigé. Nous avons étudié le comportement des EFT et des méthodes compensées avec les arrondis dirigés (par excès ou par défaut). Nous avons montré que les méthodes compensées peuvent être combinées avec l'ASD [Graillat et al., 2018a] ou avec l'arithmétique d'intervalles [Graillat and Jézéquel, 2020] qui s'appuient toutes deux sur les arrondis dirigés.

Les algorithmes 2Sum et Fast2Sum (deux algorithmes permettant de calculer l'erreur d'arrondi d'une addition) sont des éléments importants du calcul numérique. Ils sont utilisés (implicitement ou explicitement) dans de nombreux algorithmes compensés (tels que la sommation compensée ou l'évaluation polynomiale compensée). Ils sont également utilisés pour manipuler les expansions en virgule flottante. Dans [Boldo et al., 2017], nous montrons que ces algorithmes sont beaucoup plus robustes qu'on ne le croit habituellement : le résultat retourné a un sens même lorsque la fonction d'arrondi n'est pas l'arrondi au plus près, et ils n'engendrent quasiment pas de dépassement de capacité. L'algorithme 2Prod permet de calculer l'erreur d'arrondi lors d'une multiplication. Cet algorithme ne fonctionne qu'avec l'arrondi au plus près. Dans [Graillat et al., 2020], nous avons proposé un nouvel algorithme qui fonctionne aussi dans le cas des arrondis dirigés.

Nous utilisons les transformations dites "sans erreur" (EFT) pour améliorer les algorithmes classiques. Nous avons proposé un algorithme compensé pour l'évaluation des fractions rationnelles [Graillat, 2018]. Dans [Cameron and Graillat, 2022], nous avons utilisé l'algorithme de Horner compensé pour améliorer la précision de la méthode de Ehrlich-Aberth. Nous nous intéressons à l'obtention d'algorithmes précis, mais aussi efficaces sur des machines parallèles (multi-cœurs/many-cœurs/SIMD). Nous avons proposé dans [Lei et al., 2022] un algorithme parallèle de sommation très efficace permettant de calculer précisément des sommes très mal conditionnées. Dans [Graillat et al., 2017a], nous avons présenté une version parallèle de l'algorithme compensé de Horner pour l'évaluation polynomiale précise.

Dans [Iakymchuk et al., 2019], nous avons proposé une variante reproductible de la factorisation LU pour les processeurs graphiques (GPU). Pour ce faire, nous nous appuyons sur des noyaux BLAS de niveau 1 et 2 qui fournissent des résultats correctement arrondis et reproductibles pour le produit scalaire et le produit matrice-vecteur. En outre, nous élaborons une stratégie pour améliorer la précision de la résolution triangulaire par le biais du raffinement itératif. En suivant une approche ascendante, nous construisons finalement une implémentation reproductible de la factorisation de LU pour les GPU, qui prend en compte le pivotage partiel pour la stabilité et peut finalement être intégrée dans un algorithme stable et performant pour la résolution de systèmes linéaires.

Enfin, si les méthodes précédentes visent à améliorer la qualité des algorithmes en précision classique, elles ne sont pas pertinentes pour des algorithmes en précision faible, car il est moins coûteux de simplement augmenter la précision. Pourtant, la plupart des algorithmes numériques restent très vulnérables à l'utilisation de précision faible, et notamment à l'accumulation d'erreurs d'arrondi qui peut rendre le résultat de calculs à grande échelle entièrement incorrect. Pour répondre à ce défi, nous avons proposé un algorithme de sommation bloquée, dit FABsum [Blanchard et al., 2020b], qui est bien moins sensible à l'accumulation d'erreurs, et donc bien plus précis. Surtout, FABsum reste presque aussi rapide que les algorithmes de sommation classiques et est donc attractif même en précision faible.

Fiabilité des algorithmes embarqués. En collaboration avec l'équipe Inria "Toccata", nous avons travaillé sur la vérification formelle de certains algorithmes implémentés en virgule fixe. Dans le cadre d'une thèse, nous avons prouvé formellement les modèles de propagation d'erreur à travers un filtre linéaire à réponse impulsionnelle infinie (traitement du signal) implémenté en virgule fixe [Gallois-Wong et al., 2018b]. Cela ouvre la voie à la certification des codes générés pour l'implémentation en virgule fixe de filtres linéaires. De plus, dans [Boldo et al., 2020],

nous avons proposé un nouvel algorithme efficace de produit scalaire en virgule fixe correctement arrondi dans le cas des mêmes filtres linéaires. Ces travaux s'inscrivent dans une démarche plus globale qui vise à s'emparer du problème d'implémentation optimale (sous différents critères de "coût" d'implémentation) sous contrainte de précision numérique. On se recentre sur des classes d'algorithmes simples afin de formaliser complètement le problème et proposer une approche globale. Celle-ci va de la recherche des algorithmes équivalents (ré-écriture des algorithmes pour considérer des algorithmes nécessitant plus de calculs mais étant fortement moins sensibles aux erreurs d'arrondis des coefficients employés ou à la propagation des erreurs de calculs), l'évaluation de la propagation des erreurs sur la sortie (avec une approche statique, puisque la classe d'algorithme est restreinte – par opposition aux approches dynamiques, comme l'arithmétique stochastique par exemple), le problème d'allocation des largeurs et enfin la génération automatique du code. Dans ce contexte, nous nous appuyons entre autres sur la bibliothèque FiXiF⁸. Nous avons par exemple proposé une résolution du problème d'allocation des largeurs pour les algorithmes de systèmes linéaires *State-Space* dans [Hilaire et al., 2019]. Nous avons aussi mis en place une méthode pour vérifier formellement que le filtre implémenté en virgule fixe vérifie des contraintes fréquentielles données [Volkova et al., 2020].

Calculs en précision mixte. Les architectures de calcul modernes (CPU et GPU) disposent de plusieurs précisions implémentées en matériel, notamment la précision double (64 bits), simple (32 bits) et demie (16 bits). Les calculs effectués en précision faible (32 voire 16 bits) sont bien plus rapides et économes en énergie. Cependant, la plupart des applications en calcul scientifique nécessitent une solution de qualité finale équivalente à 64 bits. L'équipe travaille pour cette raison sur diverses approches *en précision mixte*, qui combinent plusieurs précisions dans le même code dans l'espoir d'obtenir une amélioration significative des performances mais sans dégrader la qualité du résultat.

L'équipe se concentre notamment sur la conception d'algorithmes en précision mixte pour l'algèbre linéaire, en particulier la résolution de systèmes linéaires de grande taille. D'un côté, nous avons amélioré les méthodes directes, fondées sur la factorisation LU de la matrice. Nous avons en effet proposé de nouvelles approches de raffinement itératif avec factorisation LU approchée et/ou en précision faible [Amestoy et al., 2022c]. Nous nous sommes aussi intéressés à l'amélioration des approches de factorisation *Block Low-Rank* (BLR) en précision mixte [Amestoy et al., 2022a]. De l'autre côté, nous avons aussi initié des travaux autour des méthodes itératives, par exemple de type Krylov. Nous avons notamment proposé des stratégies adaptatives pour le produit matrice vecteur creux et son utilisation dans les solveurs itératifs [Molina et al., 2022].

L'équipe travaille également sur le développement d'algorithmes sur matériel moderne, notamment GPU et autres accélérateurs disposant d'unités de calcul à précision faible extrêmement rapides. L'utilisation efficace et robuste de ces unités est un challenge car elles sont très spécialisées (notamment pour l'IA). Nous avons effectué une analyse d'erreur de ces unités dans un cadre générique [Blanchard et al., 2020a] qui a permis de mieux comprendre leur comportement. Cela a abouti à de nouveaux algorithmes qui exploitent ces unités pour accélérer les calculs et améliorer la qualité du résultat par rapport à une utilisation classique ou naïve de la précision faible, notamment pour la factorisation LU [Lopez and Mary, 2022]. De plus, ces unités peuvent servir à émuler une précision plus élevée par l'utilisation d'arithmétique *multiword* [Fasi et al., 2022].

Enfin, les travaux de l'équipe portent aussi sur l'auto-ajustement de la précision. L'équipe développe le logiciel PROMISE qui, à partir d'un code en C/C++ et d'une qualité numérique souhaitée sur les résultats, produit une nouvelle version du code en précision mixte [Graillat et al., 2019, Jézéquel et al., 2021]. PROMISE utilise l'algorithme de Delta-Debug [13] pour rechercher une configuration de types adéquate avec une complexité raisonnable. Une des particularités de PROMISE est le fait que les différentes configurations de types sont validées à l'aide de la bibliothèque CADNA, tandis que les autres logiciels d'auto-ajustement de la précision utilisent un résultat de référence potentiellement lui-même affecté par les erreurs d'arrondi. PROMISE a été utilisé notamment sur un code d'EDF de simulation des cœurs nucléaires. Plus récemment, PROMISE a permis d'obtenir des versions en précision mixte de codes de réseaux de neurones à partir de modèles Keras ou PyTorch [Ferro et al., 2022].

Assimilation des images et apprentissage profond. L'assimilation de données (AD) désigne un ensemble de méthodes qui permettent d'estimer l'état d'un système physique à partir d'observations partielles du système et de la connaissance de sa physique (modélisée par des EDP). Nous avons proposé une série de méthodes pour retrouver la circulation de surface de l'océan à partir de cartes satellites de température [Béréziat et al., 2017], ou bien faire la prévision opérationnelle à court terme de précipitations à partir de cartes Radar de précipitations [Zébiri et al., 2019, Zébiri et al., 2019]. Les informations dont on dispose ne sont pas parfaites (bruit dans les observations, modèles approximatifs), mais il est possible d'exploiter ces incertitudes, par exemple pour déduire des forces géophysiques [Béréziat and Herlin, 2018].

Il y a de grandes similitudes entre l'AD dans son formalisme variationnel (4DVar) et l'apprentissage statistique

8. <https://github.com/fixif/fixif>

profond (DL) dans la rétropropagation du gradient de la fonction de coût. Nous avons alors proposé dès 2020, dans le cadre de la thèse d'Arthur Filoche, un système hybride alternant les phases d'entraînement d'un modèle avec celle de l'estimation du système par AD [Filoche et al., 2020]. Ce faisant, nous déterminons à la fois l'état du système, tout en apprenant une partie de sa dynamique.

D'un côté nous avons des méthodes non supervisées (AD), de l'autre des méthodes supervisées (DL), mais en géosciences, nous ne disposons pas de vérité terrain. C'est pourquoi nous sommes convaincus que les approches AD restent d'actualité dans ce domaine. Néanmoins nous avons continué à explorer cette voie et proposé une nouvelle façon de contrôler l'optimisation 4DVar, non pas dans l'espace d'état du système, mais dans les poids d'un réseau de neurones, entraîné de manière non supervisée. Nous avons montré que le réseau, s'il est bien choisi, assure la régularisation de 4DVar [Filoche et al., 2022b]. Nous l'avons mis en œuvre pour un problème de superrésolution [Archambault et al., 2022b] et de reconstruction de champs géophysiques [Filoche and Béréziat, 2022]. Par ailleurs, cette approche dispense de connaître les matrices de covariances des termes d'erreurs, en principe difficiles à obtenir [Filoche et al., 2022a].

Animation scientifique

Des séminaires sont régulièrement organisés. Il s'agit de séminaires classiques d'une heure au cours de laquelle un collègue français ou étranger est invité et présente un résultat de recherche récent. Ces séminaires ont été organisés en ligne pendant l'année 2020. Les réunions d'équipe ne sont pas programmées régulièrement mais organisées en cas de besoin. D'autres réunions permettent de discuter de thématiques particulières ou de projets de recherche impliquant plusieurs membres de l'équipe. D'autre part, des discussions informelles nous permettent de rester informés de nos activités de recherche. Enfin les listes de diffusion à l'attention des permanents ou de toute l'équipe nous permettent des échanges réguliers sur des sujets scientifiques ou pratiques.

2 INTRODUCTION DU PORTFOLIO

Cette section identifie les éléments de portfolio présentés par l'équipe PEQUAN. Chaque élément disposant de sa propre fiche explicative, nous nous contentons ici d'en donner une liste simple :

- ▶ **Élément 1 (publication)** : Connolly, M.P., Higham, N.J., and Mary, T. (2021). Stochastic Rounding and its Probabilistic Backward Error Analysis. *SIAM Journal on Scientific Computing*, 43(1) : A566–A585.
Ces travaux répondent à une conjecture vieille de 60 ans et montrent comment l'arrondi stochastique peut améliorer la précision des calculs, ce qui explique son succès en IA et autres applications utilisant les précisions faibles.
- ▶ **Élément 2 (logiciel ou bibliothèque logicielle)** : CADNA.
La bibliothèque CADNA permet d'estimer la propagation des erreurs d'arrondi dans les codes de calcul. Elle répond au besoin crucial de validation numérique des résultats, en particulier dans les simulations à haute performance.
- ▶ **Élément 3 (publication)** : Amestoy, P., Boiteau, O., Buttari, A., Gerest, M., Jézéquel, F., L'Excellent, J.-Y., and Mary, T. (2022). Mixed Precision Low Rank Approximations and their Application to Block Low Rank LU Factorization. *IMA Journal of Numerical Analysis*.
Ces travaux décrivent une méthode innovante permettant d'utiliser plusieurs formats de précisions au sein d'un solveur à la pointe de l'état de l'art pour la résolution de systèmes linéaires creux avec approximations de rang faible. Ils présentent un fort impact applicatif avec une réduction des coûts en mémoire et en temps sur une large classe de problèmes académiques ou industriels.
- ▶ **Élément 4 (publication)** : Filoche, A., Archambault, T., Charantonis, A. A., and Béréziat, D. (2022). Statistics-free interpolation of ocean observations with deep spatio-temporal prior. In *ECML/PKDD Workshop on Machine Learning for Earth Observation and Prediction (MACLEAN)*.
Cet élément du portfolio illustre l'utilisation d'une architecture hybride assimilation de données/réseau de neurones profond pour la reconstruction de champs géophysiques épars sans connaissance sur les matrices de covariance.

3 AUTOÉVALUATION DU BILAN

3.1 Autoévaluation de l'équipe

Domaine 2. Attractivité

Référence 1. L'unité est attractive par son rayonnement scientifique et s'insère dans l'espace européen de la recherche.

Invitations dans des institutions académiques.

Différents membres de l'équipe ont été invités dans des institutions académiques, notamment au centre de recherches RIKEN (Kobe, Japon) en 2019 et 2020 et à l'université de Manchester (Royaume-Uni) en 2021. Les travaux menés avec ces deux institutions ont donné lieu à plusieurs publications communes. D'autre part, différents membres de l'équipe sont régulièrement invités à des séminaires d'une semaine à Dagstuhl (Allemagne) : sur la période évaluée, "Reliable Computation and Complexity on the Reals" en 2017 et "Resiliency in numerical algorithm design" en 2020.

Invitations dans des congrès internationaux. Du fait de leur expertise en fiabilité numérique, plusieurs membres de l'équipe ont été invités dans des congrès internationaux sur cette thématique :

- ▶ International Workshop on Numerical Verification and its Applications (INVA 2017), Okinawa (Japon) en 2017
- ▶ International symposium on Scientific Computing, Computer Arithmetic and Validated Numerics (SCAN 2020) en 2021 (online)
- ▶ International Workshop on Reliable Computing and Computer-Assisted Proofs (ReCAP 2022) en 2022 (online).

On peut mentionner aussi des invitations dans des congrès internationaux en calcul haute performance :

- ▶ Workshop on Large-scale Parallel Numerical Computing Technology, Kobe (Japon) en 2019 et 2020
- ▶ EuroHPC Summit, Green HPC session en 2021 (online)
- ▶ SimRace Conference, Numerical Methods and High Performance Computing for industrial applications, IFPEN, Rueil-Malmaison (France) en 2021.

Les interventions ont alors essentiellement porté sur la précision mixte et son impact sur la performance des codes. Les présentations de l'équipe à la conférence SimRace en 2021 ont été à l'origine de la collaboration avec IFPEN.

Conférences où des membres de l'équipe sont membres du comité de programme. Plusieurs membres de l'équipe ont été membres du comité de programme de congrès internationaux portant sur diverses thématiques :

- ▶ Calcul symbolique/numérique
 - ISSAC (International Symposium on Symbolic and Algebraic Computation) en 2019
 - SYNASC (International Symposium on Symbolic and Numeric Algorithms for Scientific Computing) en 2019
- ▶ Arithmétique des ordinateurs
 - ARITH (International Symposium on Computer Arithmetic) en 2019 et 2020
- ▶ Calcul haute performance
 - PDSEC (International Workshop on Parallel and Distributed Scientific and Engineering Computing) de 2017 à 2020
 - Euro-Par (conférence européenne dédiée au calcul parallèle et distribué) en 2022.

Organisation de conférences et de manifestations scientifiques. L'équipe a contribué à l'organisation de différentes conférences internationales portant sur la fiabilité numérique et la performance des calculs à large échelle :

- ▶ Workshop on Large-scale Parallel Numerical Computing Technology, RIKEN Center for Computational Science (R-CCS), Kobe (Japon) en 2020
- ▶ mini-symposium "Mixed Precision Algorithms for High Performance Scientific Computing", SIAM Conference on Computational Science and Engineering (SCE21) en 2021
- ▶ mini-symposium "Sustainable and Robust Linear Algebra Computations at Exascale", SIAM Conference on Applied Linear Algebra (LA21) en 2021.

D'autre part, l'équipe a organisé plusieurs conférences nationales :

- ▶ Journées nationales de calcul formel, Luminy, en 2017 et 2018
- ▶ Workshop Deep Learning & Accélération GPU à l'UPMC, Université Pierre et Marie Curie, Paris, en 2017
- ▶ 12èmes Rencontres Arithmétique de l'Informatique Mathématique (RAIM), en ligne, en 2021
- ▶ PREMIX (workshop sur la précision mixte dédié aux doctorants), Sorbonne Université, en 2022.

Enfin, l'équipe a contribué à l'organisation de manifestations scientifiques :

- ▶ école thématique CNRS "PRECIS (Précision, Reproductibilité En Calcul et Informatique Scientifique)", Fréjus, en 2017
- ▶ formation "Emerging Algorithms for Large Scale Scientific Computing" dans le cadre des formations du RFSI (Réseau Francilien en Sciences Informatiques) et de Math Innov, Paris, en 2020.

Responsabilités éditoriales dans des revues de haut niveau. L'équipe a des responsabilités éditoriales dans des revues de haut niveau (classées Q1 ou Q2 d'après SJR - Scimago Journal Rank), notamment ACM Transactions on Mathematical Software, Applied Numerical Mathematics, Numerical Algorithms.

Participation à des instances de pilotage de la recherche et d'expertise scientifique.

- ▶ Responsable d'actions à la Direction des Grands Programmes d'Investissements d'Etat (DGPIE) de l'ANR de novembre 2018 à novembre 2019
- ▶ Conseil Scientifique du Groupement de Recherche (GdR) "Calcul" du CNRS depuis janvier 2022.

Prix et distinctions. Après avoir reçu en 2018 le prix de thèse Gilles Kahn décerné par la SIF, un membre de l'équipe a obtenu 3 prix internationaux :

- ▶ Mention honorable du prix international Householder en algèbre linéaire numérique, 2020
- ▶ Leslie Fox 2nd Prize, 2021
- ▶ Prix international SIAM (Society for Industrial and Applied Mathematics) Early Career Prize décerné tous les 3 ans dans le domaine de l'algèbre linéaire, 2021.

Référence 2. L'unité est attractive par la qualité de sa politique d'accompagnement des personnels.

Sur la période évaluée, l'équipe a accueilli plusieurs post-doctorants et doctorants, deux chercheurs permanents et différents invités étrangers de renom.

- ▶ Post-doctorants, doctorants :
L'équipe accueille doctorants (5 actuellement) et post-doctorants. L'équipe a notamment accueilli : en 2020-22 un post-doctorant (docteur de l'Université d'Aachen en Allemagne) dans le cadre d'un projet "Marie Curie Fellowship" financé par l'Union Européenne, en 2021-23 une ATER (docteur de l'Université de Perpignan), en 2019-21 une ingénieure de recherche (diplômée de l'Université d'Isfahan en Iran).
- ▶ Chercheurs sur poste permanent :
Deux CR CNRS ont rejoint l'équipe, en 2019 (recrutement) et en 2022 (mutation).
- ▶ Chercheurs et enseignants-chercheurs invités de renom :
 - Thomas Cameron (Penn State Behrend, Etats-Unis) en 2022, 1 mois
 - Toshiyuki Imamura (RIKEN, Japon) en 2019 et 2022
 - Daichi Mukunoki (Tokyo Woman's Christian University puis RIKEN, Japon) en 2018, 2019 (1 mois), 2022 (2 mois)
 - Takeshi Ogita (Tokyo Woman's Christian University, Japon) en 2019
 - Sigfried Rump (Hamburg University of Technology, Allemagne) en 2018, 1 mois
 - Stepan Shakhno (Lviv University, Ukraine) en 2022, 1 mois
 - Guido Walz (Wilhelm Büchner Hochschule, Allemagne) en 2017

Les nouveaux arrivants disposent d'un ordinateur portable et éventuellement d'un ordinateur fixe. Pour effectuer des simulations numériques, d'autres ordinateurs fixes sont à la disposition des membres de l'équipe, mais aussi le cluster du LIP6, les moyens de calcul nationaux (lorsqu'un projet GENCI est en cours), le supercalculateur "fugaku" de RIKEN. L'équipe appuie les demandes de promotions ou de mobilités (en CRCT ou mutation). L'environnement de travail a permis à 2 membres de l'équipe de passer leur HDR (en 2018 et 2019).

Les doctorants participent aux séminaires de l'équipe, aux conférences ou workshops nationaux et aux conférences internationales. L'équipe a d'ailleurs co-organisé en 2022 le workshop PREMIX dédié à la précision mixte et réservé aux présentations de doctorants. Des répétitions sont organisées pour les doctorants avant leurs présentations

en séminaire ou conférence ou leur soutenance de thèse. Les doctorants sont incités à la mobilité en France ou à l'étranger. L'un d'eux a notamment été accueilli au NERSC (Nansen Environmental and Remote Sensing Center) à Bergen en Norvège pendant 2 mois en 2021. L'équipe finance au besoin la fin de thèse des doctorants dont le contrat a pris fin. Pour les post-doctorants, l'équipe organise des simulations d'auditions.

Après leur thèse, les doctorants de l'équipe partent en post-doctorat ou dans l'industrie. Deux anciens membres de l'équipe ont obtenu un poste d'enseignant-chercheur : Anastasia Volkova, doctorante dans l'équipe de 2014 à 2017, est maître de conférence à Nantes université depuis 2019 ; Roman Iakymchuk, post-doctorant dans l'équipe de 2020 à 2022, est "Associate professor" à Umea University (Suède) depuis 2022.

Référence 3. L'unité est attractive par la reconnaissance de ses succès à des appels à projets compétitifs.

Pendant les années 2017-22, l'équipe a répondu avec succès à différents appels à projets.

- ▶ Projets européens et internationaux :
 - projet autrichien COMET (COMpetence centers for Excellent Technologies) avec l'Université de Loeben (Autriche) et l'Université de Heidelberg (Allemagne), 2019-2021 ; financement d'une ingénieure de recherche pendant 2 ans
 - projet "Marie Curie Fellowship" financé par l'Union Européenne, 2020-2022 ; financement d'un post-doctorant pendant 2 ans et demi
- ▶ Programmes d'investissements nationaux :
 - PEPR NUMPEX (Numérique pour l'Exascale), 2023-28
- ▶ Projets ANR :
 - MetaLibM (générateurs de codes pour fonctions mathématiques et filtres), 2013-2017
 - FastRelax (approximation rapide et fiable), 2014-2018
 - DeRerumNatura (décider l'irrationalité et la transcendance), 2019-2024
 - NuSCAP (sûreté numérique pour les preuves assistées par ordinateur), 2021-2024
 - Interflop (plateforme d'analyse pour l'arithmétique flottante), 2021-2025
 - AVATAR (calcul haute performance pour les structures lattices), 2022-2026
- ▶ Autres appels à projets :
 - Projet DIM RFSI : Rc-Tensor (Reducing Communication in Parallel Sparse Tensor Algorithms through Mixed Precision and Data Compression Techniques) avec LISN, 2021-2024
 - Projet de la MITI (Mission pour les Initiatives Transverses et Interdisciplinaires) du CNRS : CASSIDI (CALculS fiables pour la phySIque des Deux Infinis) avec IJCLab, 2021-2024
 - Projet CNRS InFiniti ExaPol (aspects HPC en dynamique moléculaire polarisable), 2017
 - Projet Tremplin de Sorbonne Université, 2019
 - Projets financés par le LIP6 :
 - Collaboration avec l'équipe SMA du LIP6 : Accélération sur GPU de la gestion d'agendas intelligents par systèmes multi-agents, 2018 et 2019
 - Collaboration avec l'équipe POLSYS du LIP6 : GPU (Gröbner bases Performance Upwards), 2022

Référence 4. L'unité est attractive par la qualité de ses équipements et de ses compétences techniques.

Cette référence ne s'applique pas à l'équipe.

Domaine 3. Production scientifique

	2017	2018	2019	2020	2021	2022
Articles (revues)	0.92	1.07	1.40	0.90	2.25	2.00
Communications (conférences)	1.84	1.69	1.20	0.72	2.00	2.50

TABLE 2 – Publications par ETPR par an entre 2017 et 2022

PEQUAN, Évolution des publications (2017–2022)

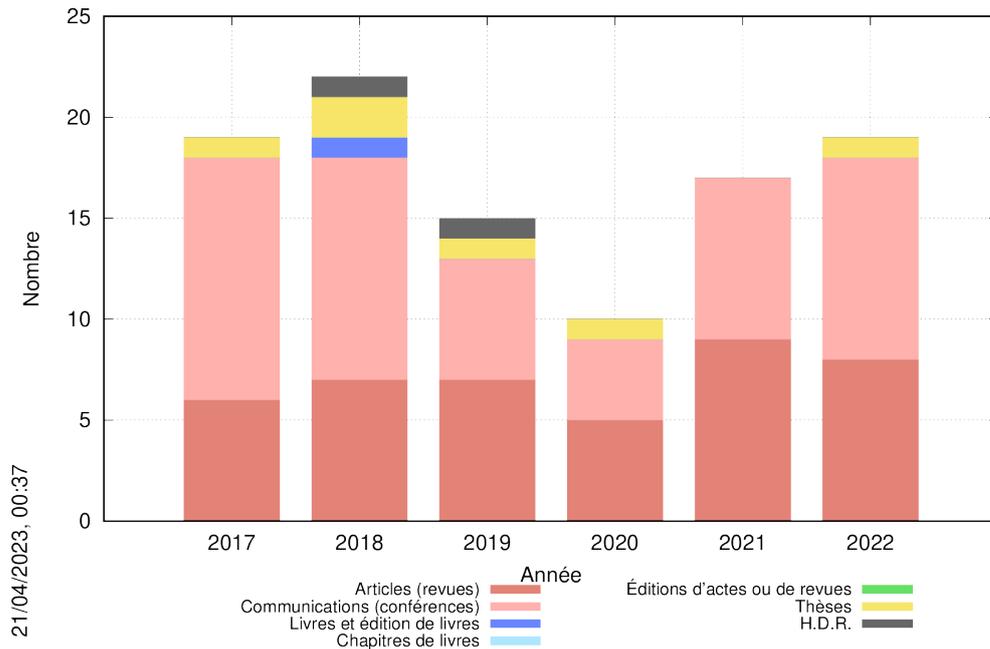


FIGURE 1 – Évolution des publications entre 2017 et 2022

Référence 1. La production scientifique de l'unité satisfait à des critères de qualité.

La production scientifique de l'équipe revêt différents aspects : conception d'algorithmes numériques, résultats théoriques (par exemple analyse fine de l'erreur commise ou des ressources de calcul consommées), implantation sur différentes architectures (par exemple architecture hybride CPU-GPU), expérimentations numériques pour des mesures effectives d'erreurs ou de performances. Les algorithmes proposés par l'équipe sont intégrés dans des logiciels libres.

Depuis ces dernières années, environ la moitié de la production scientifique de l'équipe consiste en des articles de revues internationales majoritairement Q1 telles Acta Numerica, SIAM Journal of Scientific Computing, SIAM Journal on Matrix Analysis and Applications, ACM Transactions on Mathematical Software, IMA Journal of Numerical Analysis, IEEE Transactions on Computers, Parallel Computing, International Journal of High Performance Computing Applications.

L'équipe publie avec des partenaires du meilleur niveau international tels Cleve Ashcraft (Livermore Software Technology Corporation, Etats-Unis), Nicholas Higham (University of Manchester, Royaume-Uni), Toshiyuki Imamura (RIKEN Center for Computational Science, Japon), Sherry Li (Lawrence Berkeley National Laboratory, Etats-Unis), Jean-Michel Muller (ENS Lyon, médaille d'argent 2013 du CNRS).

Référence 2. La production scientifique de l'unité est proportionnée à son potentiel de recherche et correctement répartie entre ses personnels.

La production scientifique de l'équipe a augmenté depuis la crise sanitaire et a notamment atteint en 2022 le niveau de 4.5 articles de revue ou communications en conférence par ETPR. La production scientifique est naturellement plus importante de la part des chercheurs et est impactée par les lourdes charges administratives de plusieurs enseignants-chercheurs de l'équipe. Il nous est difficile d'analyser de manière plus détaillée l'évolution de la production scientifique. Celle-ci est liée à plusieurs facteurs, notamment les différents départs de 2017 à 2022 (5 enseignants-chercheurs et 2 CR CNRS), l'arrivée de 2 CR CNRS en 2019 et 2022 et les responsabilités et mandats mentionnés dans la suite de ce document. L'arrivée des 2 CR est bien sûr positive en ce qui concerne la production scientifique et le dynamisme de l'équipe.

Les doctorants et post-doctorants participent activement à la production scientifique. Ils vont dans les conférences nationales ou internationales présenter leurs résultats et contribuent à la rédaction d'articles de revues internationales. Ils font régulièrement des présentations lors de séminaires nationaux ou internationaux. Des séminaires

sont organisés dans l'équipe avant la participation des doctorants aux conférences internationales et ces séminaires sont eux-mêmes précédés de répétitions avec leurs encadrants.

Référence 3. La production scientifique de l'unité respecte les principes de l'intégrité scientifique, de l'éthique et de la science ouverte. Elle est conforme aux directives applicables dans ce domaine.

Les manuscrits "auteur" des articles de revues ou de conférences sont accessibles via HAL. S'ils ne sont pas intégrés à un logiciel, les programmes informatiques associés sont parfois accessibles via les pages de leur auteur au LIP6. Les logiciels et données produits par l'équipe sont libres.

Domaine 4. Inscription des activités de recherche dans la société

Référence 1. L'unité se distingue par la qualité et la quantité de ses interactions avec le monde non-académique.

L'équipe collabore avec différents acteurs du monde non-académique : EDF, IFPEN, MUMPS Technologies, ONERA, Safran, Weather Measures. Plusieurs de ces collaborations ont donné lieu à des thèses CIFRE :

- ▶ EDF, avril 2015 - mars 2018 (R. Picot)
- ▶ Safran, mars 2016 - mars 2019 (C. Jeangoudoux)
- ▶ Weather Measures, mars 2017 - juin 2020 (A. Zebiri)
- ▶ EDF, novembre 2020 - 2023 (M. Gerest)
- ▶ EDF, septembre 2022 - 2025 (M. Mathevet).

Les doctorants CIFRE ayant terminé leur thèse sont maintenant en poste dans l'industrie. Ces collaborations ont notamment permis d'enrichir des logiciels tels CADNA, MUMPS ou PROMISE.

Référence 2. L'unité développe des produits à destination du monde culturel, économique et social.

L'équipe contribue fortement aux responsabilités collectives, notamment dans les structures académiques.

- ▶ Responsabilités ou mandats à Sorbonne Université :
 - Vice-président numérique de l'Université, 2016-2017
 - Directeur adjoint du LIP6, 2019-2024
 - Responsable pour l'informatique du Parcours Bi-Disciplinaire Intensif Informatique et Mathématiques, 2016-2019
 - Vice-présidente de la commission des enseignements de l'UFR d'Ingénierie, 2014-2020
 - Membre de la commission enseignements de l'UFR d'Ingénierie, 2021-
 - Responsable des contrats de moniteurs en informatique, 2014-
 - Membre du conseil de département de la licence d'informatique, 2022-
 - Membre du conseil de département du master d'informatique, 2022-
 - Co-responsable du parcours "Sécurité, Fiabilité et Performance du Numérique" (SFPN) du Master Informatique, 2013-2020
 - Co-responsable du parcours "Image" (IMA) du Master Informatique, 2016-
 - Co-responsable du parcours international "DIGIT" du Master Informatique, 2021-
 - Coordinateur local du programme VCC de l'EIT (master européen) en 2020 et 2021
 - Responsable du Master Sécurité Informatique (MSI), 2013-
 - Coordinateur de la matière "informatique" à Polytech, 2019-2021
 - Responsable de la spécialité EISE (Électronique Informatique Systèmes Embarqués) à Polytech, 2021-
- ▶ Autres responsabilités ou mandats :
 - Président du Conseil Sectoriel National du Numérique (CSNNum), 2015-2018
 - Responsable d'actions à la Direction des Grands Programmes d'Investissements d'Etat (DGPIE) de l'ANR, 2018-19
 - Directeur de l'IUT Cherbourg Manche, 2017-2019

- Coordinateur du comité stratégique sur la fusion des 3 IUT de l'université de Caen Normandie, 2017-2018
- Directeur délégué du pôle de Cherbourg-Manche de l'IUT Grand Ouest Normandie, 2019 - 2020
- Membre du Bureau du département de sciences économiques de l'Université Paris-Panthéon-Assas, 2021-

Référence 3. L'unité partage ses connaissances avec le grand public et intervient dans des débats de société.

L'équipe partage ses connaissances avec le grand public et le jeune public et participe à des actions médiation scientifique.

- ▶ Participation à la Fête de la Science, 2012-2018
- ▶ Conférence "Introduction à l'informatique" lors du stage d'un groupe de lycéens à l'Université Pierre et Marie Curie, Paris, 2017
- ▶ Conférence "Des ordinateurs, des erreurs et des hommes" donnée au LIP6 lors de l'événement "75 ans d'Informatique", 2021
- ▶ Participation à la semaine d'accueil des collégiens au LIP6, 2022

4 RÉFÉRENCES BIBLIOGRAPHIQUES EXTERNES

- [1] M. Bocquet, J. Brajard, A. Carrassi, and L. Bertino. Bayesian inference of chaotic dynamics by merging data assimilation, machine learning and expectation-maximization. *Foundations of Data Science*, 2(1), 2020.
- [2] Stefano Cherubin, Daniele Cattaneo, Michele Chiari, Antonio Di Bello, and Giovanni Agosta. TAFFO : Tuning Assistant for Floating to Fixed Point Optimization. *IEEE Embedded Systems Letters*, 12(1) :5–8, March 2020.
- [3] Wei-Fan Chiang, Mark Baranowski, Ian Briggs, Alexey Solovyev, Ganesh Gopalakrishnan, and Zvonimir Rakamarić. Rigorous Floating-Point Mixed-Precision Tuning. In *Proceedings of the 44th ACM SIGPLAN Symposium on Principles of Programming Languages*, POPL 2017, pages 300–315, New York, NY, USA, 2017. ACM.
- [4] Eva Darulova, Anastasiia Izycheva, Fariha Nasir, Fabian Ritter, Heiko Becker, and Robert Bastian. Daisy - Framework for Analysis and Optimization of Numerical Programs (Tool Paper). In *24th International Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS)*, volume 10805 LNCS, pages 270–287, Thessaloniki, Greece, 2018.
- [5] Eva Darulova and Viktor Kuncak. Towards a Compiler for Reals. *ACM Transactions on Programming Languages and Systems (TOPLAS)*, 39(2) :8 :1–8 :28, 2017.
- [6] R. Fablet, B. Chapron, L. Drumetz, E. Mémin, O. Pannekoucke, and F. Rousseau. Learning variational data assimilation models and solvers. *Modeling Earth Systems*, 13, 2021.
- [7] A. J. Geer. Learning earth system models from observations : machine learning or data assimilation ? *Phil. Trans. R. Soc. A*, 379(20200089), 2021.
- [8] Hui Guo and Cindy Rubio-González. Exploiting Community Structure for Floating-Point Precision Tuning. In *Proceedings of the 27th ACM SIGSOFT International Symposium on Software Testing and Analysis*, pages 333–343, Amsterdam Netherlands, July 2018. ACM.
- [9] M. O. Lam, T. Vanderbruggen, H. Menon, and M. Schordan. Tool Integration for Source-Level Mixed Precision. In *2019 IEEE/ACM 3rd International Workshop on Software Correctness for HPC Applications (Correctness)*, pages 27–35, 2019.
- [10] Michael O. Lam, Jeffrey K. Hollingsworth, Bronis R. de Supinski, and Matthew P. Legendre. Automatically Adapting Programs for Mixed-Precision Floating-Point Computation. In *Proceedings of the 27th International ACM Conference on International Conference on Supercomputing*, ICS '13, pages 369–378, New York, NY, USA, 2013. ACM.
- [11] Harshitha Menon, Michael O. Lam, Daniel Osei-Kuffuor, Markus Schordan, Scott Lloyd, Kathryn Mohror, and Jeffrey Hittinger. ADAPT : Algorithmic Differentiation Applied to Floating-Point Precision Tuning. In *SC18 : International Conference for High Performance Computing, Networking, Storage and Analysis*, pages 614–626, Dallas, TX, USA, November 2018. IEEE.
- [12] Cindy Rubio-González, Cuong Nguyen, Hong Diep Nguyen, James Demmel, William Kahan, Koushik Sen, David H. Bailey, Costin Iancu, and David Hough. Precimonious : Tuning Assistant for Floating-Point Precision. In *Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis*, SC'13, pages 27 :1–27 :12, New York, NY, USA, 2013. ACM.
- [13] Andreas Zeller. *Why Programs Fail*. Morgan Kaufmann, Boston, second edition, 2009.

5 RÉFÉRENCES BIBLIOGRAPHIQUES SIGNIFICATIVES DE PEQUAN

- [Amestoy et al., 2022a] Amestoy, P., Boiteau, O., Buttari, A., Gerest, M., Jézéquel, F., L'Excellent, J.-Y., and Mary, T. (2022a). Mixed Precision Low Rank Approximations and their Application to Block Low Rank LU Factorization. *IMA Journal of Numerical Analysis*.
- [Amestoy et al., 2022c] Amestoy, P., Buttari, A., Higham, N. J., L'Excellent, J.-Y., Mary, T., and Vieuble, B. (2022c). Combining sparse approximate factorizations with mixed precision iterative refinement. *ACM Transactions on Mathematical Software*.
- [Archambault et al., 2022b] Archambault, T., Filoche, A., Charantonis, A. A., and Béréziat, D. (2022b). Unsupervised Downscaling of sea surface height with Deep Image Prior. In *IA for Earth Sciences Workshop The International Conference on Learning Representations (ICLR)*, Virtual conference, United States.
- [Béréziat and Herlin, 2018] Béréziat, D. and Herlin, I. (2018). Motion and Acceleration from Image Assimilation with evolution models. *Digital Signal Processing*, 83 :45–58.
- [Béréziat et al., 2017] Béréziat, D., Herlin, I., and Lepoittevin, Y. (2017). Ensemble Kalman Filter based on the image structures. In *International Conference on Computer Vision Theory and Applications*, Porto, Portugal.
- [Blanchard et al., 2020a] Blanchard, P., Higham, N. J., Lopez, F., Mary, T., and Pranesh, S. (2020a). Mixed Precision Block Fused Multiply-Add : Error Analysis and Application to GPU Tensor Cores. *SIAM Journal on Scientific Computing*, 42(3) :C124–C141.
- [Blanchard et al., 2020b] Blanchard, P., Higham, N. J., and Mary, T. (2020b). A Class of Fast and Accurate Summation Algorithms. *SIAM Journal on Scientific Computing*.
- [Boldo et al., 2020] Boldo, S., Gallois-Wong, D., and Hilaire, T. (2020). A Correctly-Rounded Fixed-Point-Arithmetic Dot-Product Algorithm. In *ARITH 2020 - IEEE 27th Symposium on Computer Arithmetic*, pages 9–16, Portland, United States. IEEE.
- [Boldo et al., 2017] Boldo, S., Graillat, S., and Muller, J.-M. (2017). On the robustness of the 2Sum and Fast2Sum algorithms. *ACM Transactions on Mathematical Software*, 44(1).
- [Cameron and Graillat, 2022] Cameron, T. R. and Graillat, S. (2022). On a compensated Ehrlich-Aberth method for the accurate computation of all polynomial roots. *Electronic Transactions on Numerical Analysis*, 55 :401–423.
- [Connolly et al., 2021] Connolly, M. P., Higham, N. J., and Mary, T. (2021). Stochastic Rounding and its Probabilistic Backward Error Analysis. *SIAM Journal on Scientific Computing*, 43(1) :A566–A585.
- [Eberhart et al., 2018] Eberhart, P., Landreau, B., Brajard, J., Fortin, P., and Jézéquel, F. (2018). Improving CADNA Performance on GPUs. In *2018 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)*, pages 1016–1025, Vancouver, Canada. IEEE.
- [Fasi et al., 2022] Fasi, M., Higham, N. J., Lopez, F., Mary, T., and Mikaitis, M. (2022). Matrix Multiplication in Multiword Arithmetic : Error Analysis and Application to GPU Tensor Cores. *SIAM Journal on Scientific Computing*.
- [Ferro et al., 2022] Ferro, Q., Graillat, S., Hilaire, T., Jézéquel, F., and Lewandowski, B. (2022). Neural Network Precision Tuning Using Stochastic Arithmetic. In *NSV'22, 15th International Workshop on Numerical Software Verification*, Haifa, Israel.
- [Filoche et al., 2022a] Filoche, A., Archambault, T., Charantonis, A. A., and Béréziat, D. (2022a). Statistics-free interpolation of ocean observations with deep spatio-temporal prior. In *ECML/PKDD Workshop on Machine Learning for Earth Observation and Prediction (MACLEAN)*, Grenoble, France.
- [Filoche and Béréziat, 2022] Filoche, A. and Béréziat, D. (2022). Simultaneous Assimilation and Downscaling of Simulated Sea Surface Heights with Deep Image Prior. In *RFIAP (Congrès Reconnaissance des Formes, Image, Apprentissage et Perception)*, Vannes, France.
- [Filoche et al., 2022b] Filoche, A., Béréziat, D., and Charantonis, A. A. (2022b). Deep prior in variational assimilation to estimate ocean circulation without explicit regularization. In *Climate Informatics*, Asheville, NC, United States.
- [Filoche et al., 2020] Filoche, A., Brajard, J., Charantonis, A. A., and Béréziat, D. (2020). Completing physics-based models by learning hidden dynamics through data assimilation. In *NeurIPS 2020, workshop AI4Earth*, Vancouver (virtual), Canada.

- [Gallois-Wong et al., 2018b] Gallois-Wong, D., Boldo, S., and Hilaire, T. (2018b). A coq formalization of digital filters. In *Intelligent Computer Mathematics - 11th International Conference, CICM 2018, Hagenberg, Austria, August 13-17, 2018, Proceedings*, pages 87–103.
- [Graillat, 2018] Graillat, S. (2018). An accurate algorithm for evaluating rational functions. *Applied Mathematics and Computation*, 337 :494–503.
- [Graillat et al., 2017a] Graillat, S., Ibrahimy, Y., Jeangoudoux, C., and Lauter, C. (2017a). A Parallel Compensated Horner Scheme. In *CSE 2017, SIAM Conference on Computational Science and Engineering (CSE)*, Atlanta, United States.
- [Graillat and Jézéquel, 2020] Graillat, S. and Jézéquel, F. (2020). Tight interval inclusions with compensated algorithms. *IEEE Transactions on Computers*, 69(12) :1774–1783.
- [Graillat et al., 2018a] Graillat, S., Jézéquel, F., and Picot, R. (2018a). Numerical Validation of Compensated Algorithms with Stochastic Arithmetic. *Applied Mathematics and Computation*, 329 :339–363.
- [Graillat et al., 2018b] Graillat, S., Jézéquel, F., Picot, R., Févotte, F., and Lathuilière, B. (2018b). Numerical validation in quadruple precision using stochastic arithmetic. In *TNC'18. Minisymposium on Trusted Numerical Computations. International Conference on Emerging Trends in Applied Mathematics and Mechanics*, volume 8 of *Kalpa Publications in Computing*, pages 38–53, Krakow, Poland. EasyChair.
- [Graillat et al., 2019] Graillat, S., Jézéquel, F., Picot, R., Févotte, F., and Lathuilière, B. (2019). Auto-tuning for floating-point precision with Discrete Stochastic Arithmetic. *Journal of computational science*, 36 :101017.
- [Graillat et al., 2021] Graillat, S., Jézéquel, F., Queiros Martins, E., and Spyropoulos, M. (2021). Computing multiple roots of polynomials in stochastic arithmetic with Newton method and approximate GCD. In *International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC 2021)*, Timisoara, Romania.
- [Graillat et al., 2020] Graillat, S., Lefèvre, V., and Muller, J.-M. (2020). Alternative Split Functions and Dekker's Product. In *ARITH-2020 - IEEE 27th Symposium on Computer Arithmetic*, pages 1–7, Portland, United States. IEEE.
- [Higham and Mary, 2020] Higham, N. J. and Mary, T. (2020). Sharper Probabilistic Backward Error Analysis for Basic Linear Algebra Kernels with Random Data. *SIAM Journal on Scientific Computing*, 42(5) :A3427–A3446.
- [Higham and Mary, 2022] Higham, N. J. and Mary, T. (2022). Mixed precision algorithms in numerical linear algebra. *Acta Numerica*.
- [Hilaire et al., 2019] Hilaire, T., Ouzia, H., and Lopez, B. (2019). Optimal Word-Length Allocation for the Fixed-Point Implementation of Linear Filters and Controllers. In *ARITH 2019 - IEEE 26th Symposium on Computer Arithmetic*, pages 175–182, Kyoto, Japan. IEEE.
- [Iakymchuk et al., 2019] Iakymchuk, R., Graillat, S., Defour, D., and Quintana-Ortí, E. S. (2019). Hierarchical approach for deriving a reproducible unblocked LU factorization. *International Journal of High Performance Computing Applications*, page #1094342019832968.
- [Jézéquel et al., 2020] Jézéquel, F., Graillat, S., Mukunoki, D., Imamura, T., and Iakymchuk, R. (2020). Can we avoid rounding-error estimation in HPC codes and still get trustworthy results? In *NSV'20, 13th International Workshop on Numerical Software Verification*, Los Angeles, CA, United States.
- [Jézéquel et al., 2021] Jézéquel, F., Sadat Hoseininasab, S., and Hilaire, T. (2021). Numerical validation of half precision simulations. In *1st Workshop on Code Quality and Security (CQS 2021) in conjunction with World-CIST'21 (9th World Conference on Information Systems and Technologies)*, Terceira Island, Azores, Portugal.
- [Lei et al., 2022] Lei, X., Gu, T., Graillat, S., Jiang, H., and Qi, J. (2022). A Fast Parallel High-Precision Summation Algorithm Based on AccSumK. *Journal of Computational and Applied Mathematics*, 406 :113827.
- [Lopez and Mary, 2022] Lopez, F. and Mary, T. (2022). Mixed Precision LU Factorization on GPU Tensor Cores : Reducing Data Movement and Memory Footprint. *International Journal of High Performance Computing Applications*.
- [Molina et al., 2022] Molina, R., Graillat, S., Jézéquel, F., and Mary, T. (2022). Adaptive Precision Sparse Matrix-Vector Product and its application to Krylov Solvers. In *13èmes Rencontres Arithmétique de l'Informatique Mathématique (RAIM 2022)*, Nantes, France.
- [Volkova et al., 2020] Volkova, A., Hilaire, T., and Lauter, C. (2020). Arithmetic approaches for rigorous design of reliable Fixed-Point LTI filters. *IEEE Transactions on Computers*, 69(4) :489 – 504.

- 
- [Zébiri et al., 2019] Zébiri, A., Béréziat, D., Huot, E., and Herlin, I. (2019). Rain Nowcasting from Multiscale Radar Images. In *VISAPP 2019 - 14th International Conference on Computer Vision Theory and Applications*, pages 1–9, Prague, Czech Republic.
- [Zebiri et al., 2019] Zebiri, A., Béréziat, D., Huot, E., Herlin, I., and Buisson, E. (2019). Prévission multi-échelles des précipitations. In *colloque Groupe Recherche et d'Étude en Traitement du Signal et de l'Image (GRETSI)*, pages 1–4. <https://hal.sorbonne-universite.fr/hal-02164193>, <https://sites.google.com/view/gretsi2019>.

A ANNEXE — MEMBRES PERMANENTS AU 31/12/2022

La table ci dessous liste les membres permanents de l'équipe PEQUAN.

NOM	Prénom	Corps	Employeur
BÉRÉZIAT	Dominique	MCF (HDR)	Sorbonne Université
GRAILLAT	Stef	PR	Sorbonne Université
HILAIRE	Thibault	MCF	Sorbonne Université
JÉZÉQUEL	Fabienne	MCF (HDR)	Université Panthéon-Assas
JOLIVET	Pierre	CR (HDR)	CNRS
MARY	Théo	CR	CNRS
MÉNISSIER-MORAIN	Valérie	MCF	Sorbonne Université

ÉLÉMENT DE PORTFOLIO 01



Publication

1 DÉFINITION DE CET ÉLÉMENT

Titre de l'élément : Connolly, M.P., Higham, N.J., and Mary, T. (2021). Stochastic Rounding and its Probabilistic Backward Error Analysis. *SIAM Journal on Scientific Computing*, 43(1) : A566–A585.

URL de l'élément : <https://hal.science/hal-02556997>

2 MOTIVATIONS DU CHOIX DE CET ÉLÉMENT

Ces travaux répondent à une conjecture vieille de 60 ans et montrent comment l'arrondi stochastique peut améliorer la précision des calculs, ce qui explique son succès en IA et autres applications utilisant les précisions faibles.

3 PRÉSENTATION DE CET ÉLÉMENT

Les calculs en précision finie sont sujets à l'accumulation d'erreurs d'arrondi : si une opération élémentaire individuelle (addition, soustraction, multiplication, etc.) produit une erreur d'ordre u (la précision machine), un calcul plus complexe combinant plusieurs de ces opérations élémentaires cumule ces erreurs. En algèbre linéaire, la plupart des algorithmes impliquant des matrices d'ordre n (produit matrice-vecteur, résolution de système linéaire, etc.) peuvent générer une erreur au pire de l'ordre de nu . Pour des précisions faibles et/ou des problèmes de grande taille, cette borne est particulièrement préoccupante puisqu'elle suggère que le résultat est potentiellement entièrement incorrect.

Cependant, cette borne nu est une borne pire-cas : si n erreurs d'arrondi se cumulent effectivement linéairement quand elles sont de même signe, elles peuvent aussi s'annuler mutuellement si elles sont de signe opposé. Intuitivement, on peut espérer avoir environ la moitié des erreurs de chaque signe : on peut modéliser cette intuition en supposant que les erreurs d'arrondi sont des variables aléatoires d'espérance nulle, ce qui permet d'obtenir une borne d'erreur réduite de l'ordre de \sqrt{nu} . Cette ligne de pensée a été développée dès le début de l'ère digitale, dans les années 50, et a abouti à une célèbre conjecture de Wilkinson [5] déclarant que les constantes dépendant des dimensions dans les bornes d'erreur peuvent être remplacées par leur racine carrée. Cette observation est restée une conjecture, faute de preuve rigoureuse expliquant précisément sous quelles conditions cette meilleure borne tient, et avec quelle probabilité.

Dans des travaux en collaboration avec Nick Higham (Université de Manchester), nous avons prouvé la conjecture de Wilkinson. Dans une série de trois articles publiés dans *SIAM Journal of Scientific Computing* entre 2019 et 2021 [1,3,4], nous avons construit un modèle probabiliste des erreurs d'arrondis qui donne des conditions précises et suffisantes pour obtenir une borne d'erreur $\lambda\sqrt{nu}$, qui tient avec une probabilité connue $P(\lambda)$ qui se rapproche de 1 à vitesse exponentielle quand λ augmente (ainsi, la borne tient avec une très grande probabilité pour de modestes valeurs de λ). Cette nouvelle analyse d'erreur probabiliste permet aussi d'identifier certaines situations où le modèle n'est pas satisfait, et la borne pire-cas nu est atteinte.

Surtout, dans le troisième article [1], nous avons démontré un lien fort entre cette analyse probabiliste et l'arrondi stochastique. Ce mode d'arrondi, qui connaît un fort intérêt récent notamment en IA, consiste à arrondir chaque opération élémentaire de manière aléatoire ; nous nous intéressons plus particulièrement à une version de cet arrondi où la probabilité d'arrondir dans une direction ou l'autre est inversement proportionnelle à la distance de la quantité à arrondir. Or, dans [1], nous avons démontré que l'utilisation de l'arrondi stochastique force les erreurs d'arrondi à satisfaire le modèle probabiliste, et garantit donc une borne de l'ordre de \sqrt{nu} . L'arrondi stochastique est ainsi particulièrement attrayant dans les situations où la borne pire-cas est atteinte, il ainsi permet de réduire l'erreur d'un facteur \sqrt{n} et de préserver un résultat de qualité acceptable, même en précision faible. Ceci explique donc son succès dans de nombreuses applications employant une précision faible, notamment pour l'entraînement de réseaux de neurones [2].

4 RÉFÉRENCES BIBLIOGRAPHIQUES

- [1] Michael P Connolly, Nicholas J Higham, and Théo Mary. Stochastic Rounding and its Probabilistic Backward Error Analysis. *SIAM Journal on Scientific Computing*, 43(1) :A566–A585, February 2021.
- [2] Suyog Gupta, Ankur Agrawal, Kailash Gopalakrishnan, and Prithish Narayanan. Deep learning with limited numerical precision. In *Proceedings of the 32nd International Conference on International Conference on Machine Learning - Volume 37*, ICML'15, pages 1737–1746. JMLR.org, 2015.
- [3] Nicholas J. Higham and Theo Mary. A new approach to probabilistic rounding error analysis. 41(5) :A2815–A2835, 2019.
- [4] Nicholas J Higham and Théo Mary. Sharper Probabilistic Backward Error Analysis for Basic Linear Algebra Kernels with Random Data. *SIAM Journal on Scientific Computing*, 42(5) :A3427–A3446, October 2020.
- [5] J. H. Wilkinson. *Rounding Errors in Algebraic Processes*. Notes on Applied Science No. 32, Her Majesty's Stationery Office, London, 1963. Also published by Prentice-Hall, Englewood Cliffs, NJ, USA. Reprinted by Dover, New York, 1994.

ÉLÉMENT DE PORTFOLIO 02



Logiciel ou bibliothèque logicielle

1 DÉFINITION DE CET ÉLÉMENT

Titre de l'élément : Bibliothèque CADNA

URL de l'élément : <http://cadna.lip6.fr>

2 MOTIVATIONS DU CHOIX DE CET ÉLÉMENT

L'évolution des moyens de calcul est telle que la barrière de l'Exascale (10^{18} opérations en double précision par seconde) a été franchie en 2020. Or chaque opération flottante peut entraîner une erreur d'arrondi et l'accumulation de ces erreurs peut invalider les résultats calculés. La validation numérique des résultats est donc cruciale. La bibliothèque CADNA permet d'estimer la propagation d'erreur d'arrondi dans les simulations numériques. Elle peut être utilisée sur un large éventail de codes : en C/C++ ou Fortran, dans un environnement séquentiel ou parallèle, sur CPU ou GPU. Elle nécessite peu de modifications dans les codes à contrôler et permet d'estimer l'erreur commise dans tout type d'algorithme numérique. La bibliothèque CADNA a ainsi pu être utilisée avec succès pour la validation numérique de codes académiques ou industriels dans différents domaines : astrophysique, physique atomique, chimie, sciences du climat, mécanique des fluides, géophysique,...

3 PRÉSENTATION DE CET ÉLÉMENT



La bibliothèque CADNA (Control of Accuracy and Debugging for Numerical Applications) permet de contrôler les erreurs d'arrondi générées par les programmes en arithmétique à virgule flottante. CADNA implante l'Arithmétique Stochastique Discrète (ASD) [12]. Celle-ci consiste à exécuter chaque opération arithmétique plusieurs fois avec l'arrondi stochastique (par excès ou par défaut avec la même probabilité). Puis le test de Student permet, à partir de l'échantillon obtenu, d'estimer le nombre de chiffres significatifs exacts du résultat avec un intervalle de confiance à 95%. Il a été montré que 3 exécutions suffisent pour obtenir une estimation fiable de l'erreur d'arrondi commise [11]. CADNA fournit de nouveaux types (les types stochastiques) et la surcharge de tous les opérateurs et fonctions mathématiques agissant sur des variables stochastiques. CADNA nécessite donc peu de changements dans les codes à contrôler, essentiellement le remplacement des types numériques classiques par des types stochastiques. L'ASD fonctionne de manière synchrone : chaque opération est exécutée 3 fois avant de passer à la suivante. Ainsi le nombre de chiffres corrects d'un résultat peut être connu à tout moment de l'exécution du code. Ceci permet, pendant l'exécution, de détecter les instabilités numériques et aussi d'optimiser les paramètres d'une simulation numérique (pas d'intégration, seuil de convergence) de manière dynamique [8].

La bibliothèque CADNA peut être utilisée dans les codes en C/C++ et Fortran, dans les codes séquentiels ou parallèles (utilisant OpenMP et/ou MPI). Une attention particulière a été portée à l'optimisation des performances de CADNA afin que son surcoût dans les simulations numériques reste raisonnable [3]. Ces dernières années, CADNA a été étendue à la précision quadruple (128 bits) [6] et à la précision demie (16 bits) [10]. En outre, la version de CADNA qui permet de contrôler les codes sur architecture hybride CPU-GPU a bénéficié de différentes optimisations et de nouvelles fonctionnalités [4].

Par rapport à des simulations en double précision (64 bits), les simulations en précision réduite (32 ou 16 bits) ou mixte offrent l'avantage de réduire le temps d'exécution, l'empreinte mémoire et l'énergie consommée. La bibliothèque CADNA permet alors de s'assurer que ces simulations fournissent toujours des résultats de qualité numérique satisfaisante. Il est aussi possible d'utiliser le logiciel PROMISE [7] qui, à partir d'un code en C/C++ et de la précision souhaitée sur les résultats, génère une version en précision mixte de ce code. PROMISE utilise CADNA pour produire une configuration de types validée.

La bibliothèque CADNA est libre et régulièrement mise à jour à la suite des retours des utilisateurs. Le site web associé fournit différentes versions de CADNA incluant une documentation détaillée, des exemples d'utilisation, ainsi que des outils liés à CADNA : un "cadnaizer" pour l'instrumentation automatique des codes à analyser et "cadtrace" qui identifie automatiquement les instructions d'un code responsables d'instabilités numériques. Nous sommes contactés régulièrement via la liste `cadna-team@lip6.fr` pour des questions ou des demandes de fonctionnalités supplémentaires. Nous avons veillé à la facilité d'installation de CADNA, notamment en prenant en compte différents matériels, compilateurs et environnements. Les utilisateurs peuvent être des chercheurs ou ingénieurs externes inconnus de l'équipe, des partenaires nous ayant sollicités pour une collaboration, des membres (permanents ou non) de l'équipe. Les contributeurs de CADNA (une douzaine au fil des années) sont essentiellement des enseignants-chercheurs, doctorants ou stagiaires de l'équipe. CADNA a aussi bénéficié de développements et de tests effectués par l'équipe support de l'IDRIS (centre de calcul national) et de contributions de l'université de Sydney sur des aspects plus théoriques.

L'équipe contribue au projet ANR Interflop (2021-25) qui vise à proposer une plateforme unifiée d'analyse d'erreur pour l'arithmétique flottante. La bibliothèque CADNA va être intégrée dans cette plateforme, ainsi que les logiciels de validation numérique Verrou [5] et VerifiCarlo [2], eux aussi fondés sur une approche probabiliste. Ces logiciels effectuent aussi des perturbations des résultats. Contrairement à ces logiciels, CADNA utilise une approche synchrone, ce qui permet en une exécution d'obtenir à la fois la qualité numérique des résultats et un bilan des instabilités commises.

Sur la période évaluée, CADNA a généré différentes collaborations, notamment avec IFPEN (IFP Énergies Nouvelles), les laboratoires IJCLab (physique des particules) [1] et LOCEAN (sciences du climat) [4]. Dans le cadre de la collaboration de l'équipe avec RIKEN au Japon, CADNA et PROMISE ont pu être testés sur le supercalculateur "Fugaku". Des travaux menés avec RIKEN ont porté sur la réduction du coût de la validation numérique par CADNA dans les simulations haute performance [9]. Une perspective à ces travaux collaboratifs est l'optimisation des performances de noyaux de calcul en algèbre linéaire utilisant CADNA.

4 RÉFÉRENCES BIBLIOGRAPHIQUES

- [1] David Chamont, Roméo Molina, Vincent Lafage, and Fabienne Jézéquel. Investigating mixed-precision for AGATA pulse-shape analysis. In *26th International Conference on Computing in High Energy & Nuclear Physics*, Norfolk, Virginia, United States, May 2023. Jefferson Lab.
- [2] Christophe Denis, Pablo de Oliveira Castro, and Eric Petit. Verificarlo : checking floating point accuracy through Monte Carlo Arithmetic. In IEEE, editor, *2016 IEEE 23rd Symposium on Computer Arithmetic (ARITH)*, Santa Clara, United States, July 2016.
- [3] P. Eberhart, J. Brajard, P. Fortin, and F. Jézéquel. High performance numerical validation using stochastic arithmetic. *Reliable Computing*, 21 :35–52, 2015.
- [4] Pacôme Eberhart, Baptiste Landreau, Julien Brajard, Pierre Fortin, and Fabienne Jézéquel. Improving CADNA Performance on GPUs. In *2018 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)*, pages 1016–1025, Vancouver, Canada, May 2018. IEEE.
- [5] François Févotte and Bruno Lathuilière. VERROU : Assessing Floating-Point Accuracy Without Recompiling. working paper or preprint, October 2016.
- [6] S. Graillat, F. Jézéquel, R. Picot, F. Févotte, and B. Lathuilière. Numerical validation in quadruple precision using stochastic arithmetic. In Matthieu Martel, Nasrine Damouche, and Julien Alexandre Dit Sandretto, editors, *TNC'18. Trusted Numerical Computations*, volume 8 of *Kalpa Publications in Computing*, pages 38–53. EasyChair, 2018.
- [7] Stef Graillat, Fabienne Jézéquel, Romain Picot, François Févotte, and Bruno Lathuilière. Auto-Tuning for Floating-Point Precision with Discrete Stochastic Arithmetic. *Journal of Computational Science*, 36 :101017, September 2019.
- [8] Fabienne Jézéquel. Benefits of stochastic arithmetic in high performance simulations and arbitrary precision codes. In *19th international symposium on Scientific Computing, Computer Arithmetic and Validated Numerics (SCAN 2020)*, Szeged, Hungary, September 2021.

- 
- [9] Fabienne Jézéquel, Stef Graillat, Daichi Mukunoki, Toshiyuki Imamura, and Roman Iakymchuk. Can we avoid rounding-error estimation in HPC codes and still get trustworthy results? In *NSV'20, 13th International Workshop on Numerical Software Verification*, Los Angeles, CA, United States, July 2020.
- [10] F. Jézéquel, S. sadat Hoseininasab, and T. Hilaire. Numerical validation of half precision simulations. In *1st Workshop on Code Quality and Security (CQS 2021) in conjunction with WorldCIST'21 (9th World Conference on Information Systems and Technologies)*, Terceira Island, Azores, Portugal, 2021.
- [11] J. Vignes. A Stochastic Arithmetic for Reliable Scientific Computation. *Mathematics and Computers in Simulation*, 35(3) :233–261, September 1993.
- [12] J. Vignes. Discrete Stochastic Arithmetic for Validating Results of Numerical Software. *Numerical Algorithms*, 37(1–4) :377–390, December 2004.

ÉLÉMENT DE PORTFOLIO 03



Publication

1 DÉFINITION DE CET ÉLÉMENT

Titre de l'élément : Amestoy, P., Boiteau, O., Buttari, A., Gerest, M., Jézéquel, F., L'Excellent, J.-Y., and Mary, T. (2022). Mixed Precision Low Rank Approximations and their Application to Block Low Rank LU Factorization. *IMA Journal of Numerical Analysis*.

URL de l'élément : <https://hal.archives-ouvertes.fr/hal-03251738>

2 MOTIVATIONS DU CHOIX DE CET ÉLÉMENT

Ces travaux décrivent une méthode innovante permettant d'utiliser plusieurs formats de précisions au sein d'un solveur à la pointe de l'état de l'art pour la résolution de systèmes linéaires creux avec approximations de rang faible. Ils présentent un fort impact applicatif avec une réduction des coûts en mémoire et en temps sur une large classe de problèmes académiques ou industriels.

3 PRÉSENTATION DE CET ÉLÉMENT

La résolution de systèmes linéaires $Ax = b$, où A est une matrice creuse de grande taille, est un problème fondamental du calcul scientifique. Parmi les différentes méthodes pour résoudre ce type de problème, les solveurs directs (fondés sur la factorisation de la matrice A) sont très appréciés pour leur précision et leur robustesse qui leur permet de traiter tout type de problème, y compris les plus difficiles numériquement. Cependant, cette versatilité s'accompagne d'un coût calculatoire élevé en temps et en mémoire ; c'est pourquoi, en pratique, les solveurs directs de l'état de l'art exploitent des approximations numériques pour réduire leur coût. C'est notamment le cas du solveur MUMPS [2], qui exploite des approximations de rang faible par bloc (*Block Low Rank* ou BLR) pour réduire les coûts tout en préservant une précision contrôlée par l'utilisateur à travers un paramètre noté ε .

Une autre technique pour réduire les coûts de calcul consiste à utiliser des formats de précisions faibles. Cependant, passer l'intégralité du solveur en précision faible (avec une précision machine u plus grande que la précision voulue ε) dégraderait la qualité du résultat et ne permettrait donc plus l'utilisation du solveur comme une boîte noire robuste. Dans le cadre de la thèse de Matthieu Gerest (doctorant CIFRE dans l'équipe PEQUAN), en collaboration avec EDF et Mumps Technologies (en charge du développement du solveur MUMPS), nous avons développé une nouvelle méthode BLR en précision mixte, exploitant plusieurs formats de précisions au sein des représentations de rang faible par bloc. La précision de cette méthode reste contrôlable par le paramètre ε : pour chaque bloc, seuls les vecteurs singuliers associés à des valeurs singulières inférieures à ε/u sont passés en précision u , ce qui permet de maintenir une précision du même ordre ε . Par ailleurs, cette méthode peut exploiter un nombre de précisions arbitraire u_1, \dots, u_p , et est donc attrayante sur les architectures modernes disposant de nombreux formats de précision (double, simple, demie, etc.). Nous avons publié l'analyse d'erreur garantissant la robustesse de cette nouvelle méthode dans la revue *IMA Journal of Numerical Analysis* [1].

Dans le cadre de la thèse, nous avons implémenté cette méthode dans le solveur MUMPS. Une première version, qui sera mise à disposition du public dans la version MUMPS 5.6, permet de réduire les coûts mémoire du solveur. Nous avons validé le potentiel de cette version sur diverses applications, parmi lesquelles certains cas-tests industriels d'EDF mais aussi des problèmes de géophysique en collaboration avec le consortium WIND, qui a mené à une publication dans la revue *The Leading Edge* [3].

4 RÉFÉRENCES BIBLIOGRAPHIQUES

[1] Patrick Amestoy, Olivier Boiteau, Alfredo Buttari, Matthieu Gerest, Fabienne Jézéquel, Jean-Yves L'Excellent, and Théo Mary. Mixed Precision Low Rank Approximations and their Application to Block Low Rank LU Factorization. *IMA Journal of Numerical Analysis*, August 2022.

- 
- [2] Patrick R. Amestoy, Alfredo Buttari, Jean-Yves L'Excellent, and Theo Mary. Performance and scalability of the block low-rank multifrontal factorization on multicore architectures. 45(1) :2 :1–2 :26, February 2019.
 - [3] Stéphane Operto, Patrick Amestoy, Hossein Aghamiry, Stephen Beller, Alfredo Buttari, Laure Combe, Victo-rita Dolean, Matthieu Gerest, Gaoshan Guo, Pierre Jolivet, Jean-Yves L'Excellent, Frichnel Mamfoumbi, Theo Mary, Chiara Puglisi, Alessandra Ribodetti, and Pierre-Henri Tournier. Is 3D frequency-domain FWI of full-azimuth/long-offset OBN data feasible? the Gorgon-data FWI case study, 2023. Accepted for publication to special issue on FWI, The Leading Edge.

ÉLÉMENT DE PORTFOLIO 04



Publication

1 DÉFINITION DE CET ÉLÉMENT

Titre de l'élément : Filoche, A., Archambault, T., Charantonis, A. A., and Béréziat, D. (2022). Statistics-free interpolation of ocean observations with deep spatio-temporal prior. In ECML/PKDD Workshop on Machine Learning for Earth Observation and Prediction (MACLEAN).

URL de l'élément : <https://hal.sorbonne-universite.fr/hal-03765735>

2 MOTIVATIONS DU CHOIX DE CET ÉLÉMENT

L'assimilation de données utilise l'ensemble des informations disponibles pour reconstruire l'état d'un système physique tel que l'atmosphère ou l'océan. En particulier, elle nécessite la connaissance approximative de l'état du système, appelée ébauche, et de la covariance associée aux erreurs induites par cette ébauche (on fait classiquement une hypothèse de loi gaussienne). Cette matrice de covariance, notée B dans la suite, n'est pas toujours connue en pratique mais peut s'obtenir par analyse statistique d'un historique des états passés du système étudié.

Nous proposons une approche alternative aux méthodes utilisées en assimilation de données qui permet de lever l'hypothèse gaussienne faite sur l'erreur de l'ébauche. Cette approche utilise un réseau de neurones profond dont l'architecture permet d'encoder les corrélations spatiales du vecteur d'état décrivant le système. Dans un régime de changement climatique, l'analyse statistique de l'historique perd en partie de sa pertinence, notre approche permettrait d'y répondre.

3 PRÉSENTATION DE CET ÉLÉMENT

En notant X l'état d'un système, Y les observations partielles de ce système, \mathbf{H} l'opérateur d'observation (il projette l'état vers l'espace des observations), X_b une ébauche de l'état du système, l'approche standard 3DVar propose de reconstruire X en minimisant la fonction objectif :

$$J(X) = \|Y - \mathbf{H}X\|_R^2 + \|X - X_b\|_B^2 \quad (1)$$

où l'on note $\|Z\|_A^2 = \langle Z, A^{-1}Z \rangle$, c'est-à-dire que l'on suppose que les deux termes d'erreur dans J suivent une loi gaussienne dans un formalisme bayésien. Si le bruit sur les observations est connu, il est décrit par la matrice de covariance R , celui sur le terme d'ébauche ne l'est pas. En pratique la matrice B est conçue par des experts ou par analyse statistique.

Nous avons proposé de contrôler la minimisation de J dans l'espace des paramètres d'un réseau de neurones [1] correctement choisi. Ainsi la fonction objectif devient :

$$J(\theta) = \|Y - \mathbf{H} \circ g_\theta(z)\|_R^2 \quad \text{s.t.} \quad X = g_\theta(z)$$

et est illustrée par le diagramme de la figure 1. Ainsi, le terme d'ébauche disparaît, et le réseau g_θ est alors

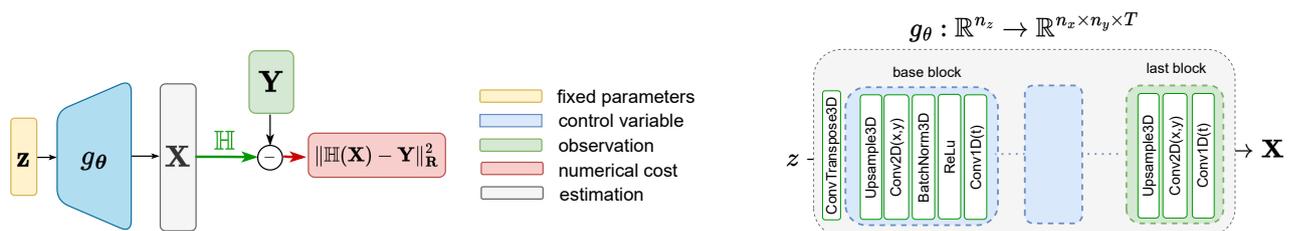


FIGURE 1 – Diagramme de Deep Prior 3DVar

entraîné sur un seul exemple : la donnée que l'on souhaite reconstruire, au contraire des approches habituelles en apprentissage statistique qui sont entraînées sur un corpus d'exemples.

Illustrons cet algorithme sur le problème de reconstruction d'une carte de hauteur d'eau à partir d'observations très partielles qui sont des traces de deux satellites (NADIR et SWOT), tel que montré par la figure 2. La hauteur d'eau est l'une des variables décrivant l'état de l'océan. Avoir une observation de ce champ sur l'ensemble de la grille est nécessaire pour alimenter les modèles océanographiques. Ici on dispose d'une vérité terrain car la méthode est évaluée sur la sortie d'un modèle de simulation numérique de circulation océanographique (NEMO, couramment utilisé par la communauté des océanographes) et une simulation des traces satellites réalisée par un algorithme de l'état de l'art.

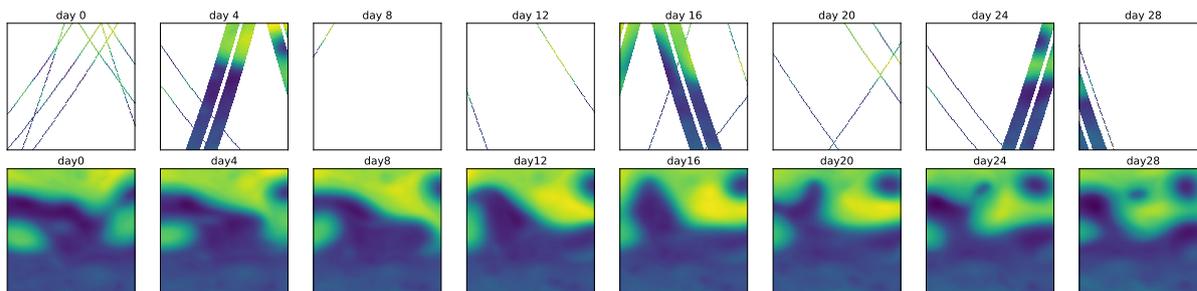


FIGURE 2 – Traces satellites NADIR et SWOT de hauteur d'eau et vérité terrain associée

La méthode de référence est DUACS [4], il s'agit d'une interpolation optimale utilisant 25 années d'observation et d'analyse pour estimer B qui minimise la fonctionnelle de l'équation 1.

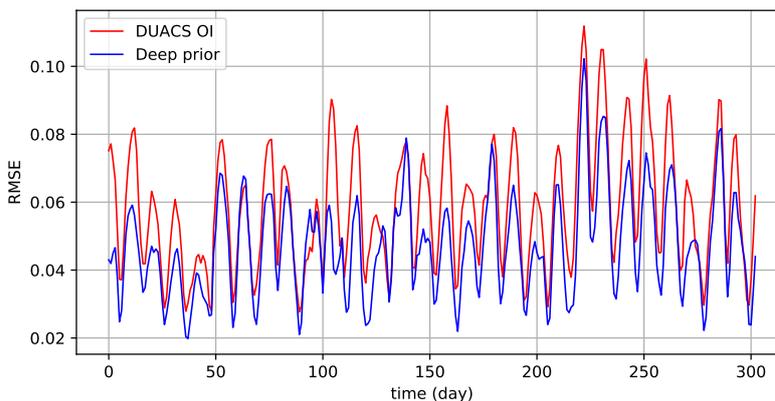


FIGURE 3 – Comparaison entre DUACS et Deep Prior 3DVar sur une année (métrique : erreur quadratique moyenne).

Comme on peut le voir en figure 3, notre méthode bat quasi-systématiquement DUACS sans information sur la matrice B . En conclusion, un réseau génératif profond peut agir comme une matrice d'ébauche B dans 3DVar. Le principe est que le réseau utilise des convolutions 3D pour capturer des corrélations spatio-temporelles compatibles avec la réalité physique. Cette approche a également été utilisée pour retrouver des cartes de circulation qui ne sont jamais observées en observant uniquement des cartes de hauteurs d'eau incomplètes en temps [3], ou pour faire simultanément la reconstruction de la circulation et/ou augmenter significativement (facteur 8 par 8) la résolution spatiale des cartes de hauteurs d'eau [2].

4 RÉFÉRENCES BIBLIOGRAPHIQUES

- [1] Arthur Filoche, Théo Archambault, Anastase Alexandre Charantonis, and Dominique Béréziat. Statistics-free interpolation of ocean observations with deep spatio-temporal prior. In *ECML/PKDD Workshop on Machine Learning for Earth Observation and Prediction (MACLEAN)*, September 2022.
- [2] Arthur Filoche and Dominique Béréziat. Simultaneous Assimilation and Downscaling of Simulated Sea Surface Heights with Deep Image Prior. In *RFIAP*, July 2022.

- 
- [3] Arthur Filoche, Dominique Béréziat, and Anastase Alexandre Charantonis. Deep prior in variational assimilation to estimate ocean circulation without explicit regularization. In *Climate Informatics*, Asheville, NC, United States, May 2022.
 - [4] G. Taburet, A. Sanchez-Roman, M. Ballarotta, M.-I. Pujol, J.-F. Legeais, F. Fournier, Y. Faugere, and G. Dibarboure. DUACS DT2018 : 25 years of reprocessed sea level altimetry products. *Ocean Science*, 15(5) :1207–1224, 2019.