

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.