Rounding Error Analysis of Fast Fourier Transforms for Deep Learning Applications

Keywords: Floating-Point Arithmetic, Rounding Error Analysis, Fast Fourier Transform, Deep Learning, Neural Networks, Stochastic Rounding, Probabilistic Error Analysis, Mixed Precision Arithmetic

Supervisors: Nicolas Brisebarre[1], Claude-Pierre Jeannerod[2], Theo Mary[3], Jean-Michel Muller[4] and Joris Picot[5]

Context: With the rise of computationally intensive deep learning applications such as large language models (LLM), reducing the power consumption, memory footprint, and training and inference time of neural network computations has become one of the most pressing and important challenges. Two very different classes of methods have been explored and have encountered much success: first, low precision arithmetics (e.g., 16-bit or lower), nowadays ubiquitous in deep learning, are used to quantize the network weights to smaller sizes, while accuracy is retained with the use of stochastic rounding or mixed precision arithmetic. Second, the matrices associated with the network layers are compressed by exploiting structured representations, such as sparse, low-rank, or butterfly representations. Butterfly matrices in particular have attracted growing interest due to their extreme sparsity and strong expressivity, as they are a fundamental tool in many fast linear transforms such as the Fast Fourier Transform (FFT). While both techniques (quantization and compression) have been extensively studied independently, their combined use remains largely unexplored, and raises several important questions, both theoretical and practical.

Objectives: This internship will investigate the accuracy of the FFT and related operations with structured matrices in the presence of rounding errors, and in settings typical of deep learning applications. The FFT was introduced in 1965 by Cooley and Tukey in its modern form [4, 5, 10]. There is a large literature on the error analysis of the FFT (see [13, 7, 11, 12]), most authors bounding the relative mean-square error. For some of our applications, we also need bounds in terms of the infinity norm; in particular, such an analysis was made by Henrici [6] and, recently, an improved bound was given in [2] together with the construction of "bad input cases", for which the attained error is close to this new bound.

While providing suitable worst-case models, such error analyses tend to be highly pessimistic for average case situations. Indeed these analyses do not exploit specific features of the hardware or the application that lead to small errors in practice. In particular, deep learning computations commonly employ mixed precision matrix-multiply accumulate operations [1] and/or stochastic rounding [3], both of which can reduce the error accumulation. Moreover, in the recent years probabilistic analyses have been proposed to model the somewhat random nature of both rounding errors [8] and the matrix coefficients [9] in typical neural network applications. However, these more refined analyses have been focused on dense (unstructured) linear algebra computations. The goal of this internship is therefore to extend these analyses to structured computations, including the FFT, in order to improve the current worst-case bounds. The improvement of the error bounds will be experimentally validated by performing an empirical study of the average accuracy of structured matrix products in deep learning applications.

Practical details

- The internship will take place at LIP laboratory of ENS de Lyon.
- The intern will receive indemnities if her/his status allows.
- Depending on the results of the internship, an opportunity to continue with a PhD thesis could be available.

[1] https://perso.ens-lyon.fr/nicolas.brisebarre/
[2] https://perso.ens-lyon.fr/claude-pierre.jeannerod/
[3] https://perso.lip6.fr/Theo.Mary/
[4] https://perso.ens-lyon.fr/jean-michel.muller/
[5] https://perso.ens-lyon.fr/joris.picot/

References

[1] P. Blanchard, N. J. Higham, F. Lopez, T. Mary, and S. Pranesh. Mixed Precision Block Fused Multiply-Add: Error Analysis and Application to GPU Tensor Cores. *SIAM J. Sci. Comput.*, 42(3):C124–C141, 2020.

[2] N. Brisebarre, M. Joldeş, J.-M. Muller, A.-M. Naneş, and J. Picot. Error analysis of some operations involved in the Cooley-Tukey fast Fourier transform. *ACM Trans. Math. Software*, 46(2):Art. 11, 27, 2020.

[3] M. P. Connolly, N. J. Higham, and T. Mary. Stochastic Rounding and its Probabilistic Backward Error Analysis. *SIAM J. Sci. Comput.*, 43(1):A566–A585, 2021.

[4] J. Cooley and J. Tukey. An algorithm for the machine calculation of complex Fourier series. *Math. Comp.*, 19(90):297–301, 1965.

[5] P. Duhamel and M. Vetterli. Fast Fourier transforms: a tutorial review and a state of the art. *Signal Process.*, 19:259–299, 1990.

[6] P. Henrici. *Applied and Computational Complex Analysis, Vol. 3*. Wiley, New York, 1986.

[7] N. J. Higham. *Accuracy and Stability of Numerical Algorithms*. SIAM, Philadelphia, PA, 2nd edition, 2002.

[8] N. J. Higham and T. Mary. A New Approach to Probabilistic Rounding Error Analysis. *SIAM J. Sci. Comput.*, 41(5):A2815–A2835, 2019.

[9] N. J. Higham and T. Mary. Sharper Probabilistic Backward Error Analysis for Basic Linear Algebra Kernels with Random Data. *SIAM J. Sci. Comput.*, 42(5):A3427–A3446, 2020.

[10] C. V. Loan. *Computational Frameworks for the Fast Fourier Transform*. Frontiers in Applied Mathematics. SIAM, 1992.

[11] C. Percival. Rapid multiplication modulo the sum and difference of highly composite numbers. *Math. Comp.*, 72:387–395, 2002.

[12] G. Plonka, D. Potts, G. Steidl, and M. Tasche. *Numerical Fourier Analysis*. Appl. Numer. Harm. Anal. Birkhäuser, 2018.

[13] G. Ramos. Roundoff error analysis of the fast Fourier transform. *Math. Comp.*, 25(116):757–768, 1971.