Postdoctoral Position: Advanced Algorithms for Large-Scale Training

Job Offer: A fully funded postdoctoral position (1-2 years) available immediately.
Advisor: Dr. Edouard Oyallon (CNRS, Sorbonne University), `edouard.oyallon@cnrs.fr`
Main Location: ISIR, Sorbonne University: Jussieu campus, centrally located in Paris.
Note: Possible collaboration with MILA under Pr. Eugene Belilovsky.
Application: Submit a CV, research statement, and contact details for at least two referees.

Context: Large-scale deep learning training faces computational challenges such as inefficient Model FLOPS Utilization (MFU), communication bottlenecks, and scalability limitations. This postdoctoral position aims to push MFU toward its theoretical limits by designing more parallelizable algorithms for exascale (and decentralized) multi-GPU training.
This research (anchored in : [1, 2, 3, 4, 5]) can be integrated into an existing project, allowing the candidate to contribute meaningfully while expanding their expertise in large-scale optimization. The focus includes optimizing communication, developing asynchronous execution frameworks, improving memory efficiency, and advancing model parallelism. Both applied and theoretical contributions are valued, ensuring practical impact on large-scale training. The position is flexible and can adapt to the candidate's background and research interests, offering collaboration opportunities and a chance to deepen expertise in scalable deep learning.

References

[1] A. Nabli, L. Fournier, P. Erbacher, L. Serrano, E. Belilovsky, and E. Oyallon. Acco: Accumulate while you communicate, hiding communications in distributed llm training, 2024. Preprint.

[2] A. Nabli, E. Belilovsky, and E. Oyallon. A2cid2: Accelerating asynchronous communication in decentralized deep learning. In *NeurIPS*, 2023.

[3] S. Rivaud, L. Fournier, T. Pumir, E. Belilovsky, M. Eickenberg, and E. Oyallon. Petra: Parallel end-to-end training with reversible architectures, 2025. ICLR.

[4] A. Nabli and E. Oyallon. Dadao: Decoupled accelerated decentralized asynchronous optimization. In *ICML*, 2023.

[5] B. Thérien, C.-É. Joseph, B. Knyazev, E. Oyallon, I. Rish, and E. Belilovsky. Lo: Compute-efficient meta-generalization of learned optimizers, 2024. Preprint.