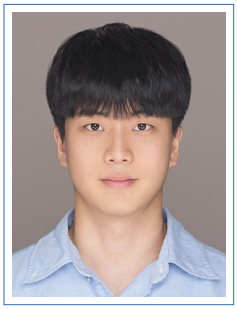
Chaewon Kim

Seoul National University
Dept. of Computer Science & Engineering
Seoul, Republic of Korea
(+82) 10 5095 2407
chaewon@aces.snu.ac.kr



Research Summary

I am a systems researcher focused on high-performance GPU computing and numerical algorithms. I am motivated by mathematical reasoning and algorithmic problem-solving, which guide my work on FFTs, sparse solvers, and accelerator-optimized computation. My research aims to develop hardware-aware methods that reduce data movement, improve locality, and deliver predictable high utilization on modern accelerators. Ultimately, I hope to build general techniques that allow future numerical and sparse workloads to run efficiently across diverse architectures.

Education

2013 Mar – 2016 Feb — **Seoul Science High School**

2016 Mar – 2022 Aug — **B.S. in Computer Science & Engineering**, *Seoul National University*
Minor in Mathematics

2022 Sep – 2025 Aug — **M.S. in Computer Science & Engineering**, *Seoul National University*

Research Experience

Sep 2022 – Present — **Graduate Researcher, Thunder Research Lab (Advisor: Prof. Jaejin Lee)**, *Seoul National University*

- **Tensor-Core FFT kernels (Mar. 2025 – Present)** — developed a Tensor Core–accelerated FFT kernel achieving over 3× speedup relative to cuFFTDx. Introduced a permutation-fused data layout that eliminates shared-memory–bound MMA stages and reduces FLOP overhead, enabling >70% Tensor Core utilization on A100/H100 GPUs. Conceived and implemented all CUDA/Tensor Core kernels, performance modeling, and evaluation.

- **Sparse linear solver optimization (Sep. 2023 – Jun. 2025)** — designed a GPU-aware nested dissection ordering and sparsity-region partitioning strategy to improve locality and reduce communication overhead in sparse LU factorization. Implemented the full heterogeneous CPU–GPU solver pipeline (data layout, kernel fusion, multi-GPU scaling) and validated substantial improvements on large scientific workloads (ICS 2025).

- **Large-scale LLM systems research (May 2022 – Mar. 2023)** — contributed to a nationwide foundation-model project covering data infrastructure, GPU optimization, and efficient training. Built large-scale crawling and preprocessing pipelines, and developed a GPU-accelerated LSH-based deduplication framework (arXiv:2501.01046). Explored resource-efficient training techniques including quantization, small-model training and merging, and Mixture-of-Experts (MoE); conducted performance/energy profiling using CUPTI and Nsight.
- **Quantum circuit optimization (Nov. 2024 – Feb. 2025)** — explored memory-aware partitioning of large quantum circuits to improve locality and reduce memory traffic. Implemented CUDA/C++ kernels that increase data reuse and reduce memory footprint, improving execution predictability.

Mar 2021 – Aug 2022 **Undergraduate Researcher, Thunder Research Lab (Advisor: Prof. Jaejin Lee)**, *Seoul National University*
- Developed GPU kernels and simulations in OpenCL/CUDA/PyTorch; investigated practical quantization strategies for ML workloads.

Sep 2020 – Dec 2020 **UROP Researcher (Advisor: Prof. Kunsoo Park)**, *Seoul National University*
- Improved the DAF graph matching algorithm by identifying and analyzing computational bottlenecks on large real-world graphs.

Feb 2020 – Aug 2020 **Research Intern, CryptoLab (Advisor: Prof. Jung Hee Cheon)**, *Seoul National University*
- Optimized the CKKS homomorphic encryption (HEAAN) library by improving polynomial operations and accelerating key routines via multi-threading and GPU parallelization.
- Developed the *Efficient Homomorphic Max Algorithm for Multivariables*, proposing a division-free polynomial formulation for encrypted comparison and analyzing its convergence and stability properties.

Publications & Preprints

1. **C. Kim**, S. Bae, S. Lee, J. Lee, J. Kim, D. Kim, K. Ahn, and J. Lee. *Accelerating Fast Fourier Transform with Tensor Cores.* Submitted for publication (PLDI, under review), 2026.

2. S. Bae, J. Lee, **C. Kim**, S. Lee, and J. Lee. *Accelerating 2D Convolution using cuFFTDx.* Korea Computer Congress (KCC), 2025.

3. **C. Kim**, J. Lee, J. Kim, D. Kim, K. Ahn, H. U. Cho, S. Baek, and J. Lee. *SnuSOLVER: Optimizing Sparse Direct Solvers for Heterogeneous Systems.* Proceedings of ICS 2025.

4. Y. Son, **C. Kim**, and J. Lee. *FED: Fast and Efficient Dataset Deduplication Framework with GPU Acceleration.* arXiv:2501.01046, 2025.

5. S. Min, W. Shin, **C. Kim**, and K. Park. *New Adaptive Matching Order and Performance Comparison for Subgraph Matching Problem.* KCC Best Paper Award, 2021.

6. J. H. Cheon, W. Cho, S. Hong, and **C. Kim**. *Efficient Homomorphic Max Algorithm for Multivariables.* Preprint, 2020.

Awards & Honors

2016–2022 **Korea Student Aid Foundation** — Presidential Science Scholarship

2021 **ACM-ICPC Asia Regional (Seoul)** — 4th Place

2020 **ACM-ICPC Asia Regional (Seoul)** — 4th Place

2019 Korean Mathematical Competition for University Students — Silver Prize

2019 University Collegiate Programming Contest (UCPC) — 3rd Place

2018 Samsung Collegiate Programming Cup — 5th Place

2018 Simon Marais Mathematics Competition — 7th Place

2015 **International Mathematical Olympiad** — Bronze Medal

2015 **National Team Selection** — Selected as one of six representatives

2014 Korea Mathematical Olympiad — Gold Medal

Research Interests

- High-performance GPU and accelerator systems; parallel and heterogeneous computing
- Algorithm–hardware co-design; compiler IR, kernel generation, and runtime optimization
- Sparse computation and sparsity-driven algorithms (irregular sparsity patterns, graph-structured sparsity, block/hierarchical sparsity)
- Numerical algorithms (FFT, linear algebra) and data-movement–efficient computation
- Distributed and multi-GPU systems; communication-efficient parallel algorithms

Other Experiences

Military Service: Republic of Korea Air Force — Special Intelligence (2018 Mar – 2020 Feb)

Teaching: Teaching Assistant, Korean Mathematical Olympiad Winter Camp (2017, 2018)
Teaching Assistant, Accelerator Programming School (Winter 2022 – Summer 2025)
Teaching Assistant, High Performance Parallel Computing (Fall 2022 – Spring 2025)
Teaching Assistant, Algorithms (Spring 2025)

Languages: Korean (native); English (fluent)

Technical Skills: C/C++, Python; CUDA and GPU kernel development; performance modeling and optimization; parallel algorithms; PyTorch and ML systems; MPI/NCCL; LLVM; Nsight/CUPTI profiling.