

# ZENG, ZHICHEN

• ☎: 607-327-2454

• ✉:zhichen.zengustc@gmail.com

• Home: zhichenzzz.github.io

## EDUCATION BACKGROUND

### University of Science and Technology of China (USTC)

*Bachelor's Degree of Science*

09/2020 – 07/2024 (Expected)

*Yan Jici Talent Program in Physics, Major: Physical Electronics*

GPA: **4.08/4.30**, Ranking: **1/32** (Major), **3/183** (School)

**Core Courses:** *Computer Architecture (91), Principles of Microcomputer (91), Digital Logical Circuit (95), Electronic Circuits (95), Computational Physics A (95), Thermodynamics and Statistical Physics (100), Electromagnetism (100), Quantum Mechanics (95), Quantum Computing and Machine Learning (91), Computer Programming A (91), High-Level Digital Design Automation (audited, Cornell ECE 6775)*

## PUBLICATIONS

- Hongzheng Chen\*, Niansong Zhang\*, Shaojie Xiang, **Zhichen Zeng**, Mengjia Dai, Zhiru Zhang. “Allo: A Programming Model for Composable Accelerator Design”, (*PLDI 2024, under review*).
- Qizhe Wu, Yuchen Gui, **Zhichen Zeng**, Xiaotian Wang, Huawen Liang, Linfeng Tao, Letian Zhao, Zhaoxi Zeng, Xi Jin. “EN-Tensor: Advancing TensorCores Performance through Encoder-Based Methodology”, (*DAC 2024, under review*).
- Zhengyu Han\*, Mengjia Dai\*, **Zhichen Zeng\***, Chunhui Ye, Rucheng Dai, Zhongping Wang, Xiaoyu Sun, Zengming Zhang. “Highly Stable Photoelectric Detector Using Lead-Free Double Perovskite Cs<sub>2</sub>AgBiCl<sub>6</sub> with Fast Response”, *Journal of Materials Chemistry C (JMCC, under review)*.

## RESEARCH INTERESTS

- Accelerator architectures (especially FPGA), Programming Abstraction on Heterogeneous Hardware, and Domain-specific Compiler.

## RESEARCH EXPERIENCE

### Allo: A Programming Model for Composable Accelerator Design

07/2023 – Present

Research Assistant, Advisor: **Prof. Zhiru Zhang, Computer Systems Laboratory**, Cornell University NY, USA

- Contributed to Allo, a MLIR-based programming infrastructure with decoupled customizations that target CPU, FPGA, and AI Engines.
- Proposed new programming interface and frontend framework to support advanced Python-based features (based on LLVM Linalg dialect operations).
- Developed a model tracer to compile PyTorch models (GPT, BERT, CNNs) to MLIR on different hardware devices.
- Optimized KV cache on GPT module and realized end-to-end GPT computation on FPGA with **1.76x** times faster than the A100 GPU and **5.6x** times higher energy efficiency.
- Outperformed ScaleHLS for all test cases in PolyBench with less latency and resource usage.

### EN-Tensor: Advancing Tensor Cores Performance through Encoder-Based Methodology

03/2023– 07/2023

Research Assistant, Advisor: **Prof. Xi Jin, SoC Design Lab**, USTC

Hefei, China

- Proposed an Encoder-TensorCore architecture that extracts encoders from all the multipliers within systolic array units to reduce redundancy of encoding.
- Validate the feasibility of the novel radix-4 encoding and expand the radix to arbitrary integer.
- Support both INT and FLOAT with N-bits matrix multiplication with higher performance.
- Reduced logic area and power consumption by 10%-20% on mainstream NPU architectures.

## Useful Research Experiences

*Done in USTC, Hefei, China*

### ➤ **Highly Stable Photoelectric Detector Using Lead-Free Double Perovskite $\text{Cs}_2\text{AgBiCl}_6$ with Fast Response**

*Advisor: Prof. Zengming Zhang, Physics Experimental Center, USTC*

- (1) Designed a highly stable photoelectric detector which has a fast response of light (~ms).
- (2) Explored the photoelectric properties of the detector which behaves excellent in high-temperature (473K) and vacuum ( $8 \times 10^{-5}\text{Pa}$ ) conditions.

### ➤ **Used Monte-Carlo and Metropolis Sampling Methods to Simulate Several Physical Problems**

*Advisor: Prof. Zejun Ding, Computational Physics, USTC*

- (1) Brownian motion of particles in an electric field, Diffusion-limited aggregation (DLA), and Dielectric breakdown model (DBM)
- (2) Percolation calculation, and 2D-Ising module

### ➤ **Designed Multi-Cycle Baseline CPU with five stages pipeline by Using Verilog HDL**

## SCHOLARSHIPS

- USTC Fellowship Top 1% for Research, Chinese Academy of Sciences & USTC 08/2023
- Outstanding Student Scholarship, Gold (1%), USTC 10/2022
- Scholarship for Talent Program in Basic Disciplines, Class A, USTC 09/2022, 2021
- Yilin Chen Foundation Scholarship (2%) 09/2021

## REWARDS/COMPETITIONS

- USTC Physics Research, grand prize (1%) 04/2023
- Atomic Physics Course Thesis Competition, Second Prize, USTC 01/2021
- Mathematics Competition of Chinese College Students, First Prize, the Ministry of Education 10/2021

## EXTRACURRICULAR ACTIVITY

- Teaching Assistant
  - *Electrodynamics*, by Prof. Qing jia, Spring 2023
- Member of the Association of Basketball of USTC 10/2020-07/2022

## SKILLS/STANDARDIZED TESTS

- Computer skills: C/C++, Python, PyTorch, MLIR, Verilog HDL, Vitis HLS, Git, MATLAB, Mathematica, LATEX, Origin, Linux (Ubuntu)
- TOFEL: 99 (R:26, L:23, S:23, W:27)