Amin Ghasemazar

+1(604)728-2171
aming@ece.ubc.ca
ece.ubc.ca/~aming
linkedin.com/in/amin-ghasemazar/

ABOUT ME

I am a last year Ph.D. Candidate working in computer systems. For my research I mainly focused on (1) Hardware acceleration for training sparse neural networks, (2) In-hardware compression of on-chip data, and (3) Approximate computing. I am experienced in software programming and hardware design and I have published papers in top-tier venues including *ASPLOS*, *MICRO* and *DATE*.

EDUCATION

Candidate for Doctor of Philosophy, Electrical and Computer Engineering. *[Sept. 2015 – **expected 2021**]*
The University of British Columbia (UBC)
Thesis: Efficient In-Hardware Compression of On-Chip Data

Master of Science, Computer Engineering. *[Sept. 2012– Aug. 2015]*
University of Tehran (UT)
Thesis: Improving Efficiency of Extensible Processors by Using Approximate Custom Instructions

Bachelor of Science, Computer Engineering, Hardware Systems. *[Sept. 2008– Aug. 2012]*
Iran University of Science and Technology (IUST)
Capstone: Configurable high-power wireless mesh network using embedded Linux machines

EXPERIENCES

Research Internship – *MediaTek Inc., Boston, MA [Aug. 2017 – Nov. 2017]*
- Working in DSP architecture team, proposed and investigated HW/SW methodologies leveraging low-precision computing for efficient storage and wireless communications in modems. Results showed significant benefits with negligible quality loss.

Creative Destruction Lab (CDL) / Consultation – *Vancouver, BC [Sept. 2019–Apr. 2020]*
- Support CDL-West companies with their technical needs, market research, competitive analysis and customer development in the *CDL's educational* program.

Research Assistant – *UBC, Vancouver, BC [Sept. 2015 – present]*
- **AI Acceleration:** Designed deep learning accelerations leveraging data compression and sparse training.
- **Data compression:** Developed data compression mechanisms to increase effective cache capacity.
- **Approximation:** Proposed an accurate error estimation mechanism for approximate circuits.

PROJECTS

AI / Domain-Specific Accelerators
- ***Procrustes sparse training:*** Designed a hardware accelerator for training sparse neural networks in a group of 3. Identified challenges including load imbalance and dataflow and proposed an accelerator architecture to enable sparse training from scratch using a hardware-friendly sparse training algorithm. Work Accepted in MICRO 2020.
- ***Accelerated FP-ALU:*** Proposed a new complex floating point ALU in pipelined MIMO fashion and implemented it in RTL level. In an FPGA implementation, the design resulted in up to 12× speed up over state-of-the-art designs. Work published in VLSID 2015.

On-Chip Data Compression Techniques
- ***Thesaurus compression:*** Developed a novel cache compression scheme based on hardware-level on-line cacheline clustering. The online clustering mechanism leverages the locality-sensitive hashing (LSH). Work published in *ASPLOS 2020*.
- ***2D cache compression:*** Proposed a compression scheme that leverages both intra-block and inter-block (across cachelines) compression. Work published in *DATE 2020*.

Deep Neural Networks
- ***Extended-PyTorch:*** Extended PyTorch framework to support pruning, quantization and approximate neural networks. To achieve these, the framework intervenes and captures values of the network being trained, modifies it as necessary and feeds it back to the training pipeline. It also supports various types of regularization on weights and activations to sparsity DNNs. Part of ongoing research work.

- **TrainMe:** Implemented a neural network training management infrastructure for Identifying the training tradeoffs in order for optimizing hyperparameters on multi-GPU setups. Part of MICRO 2020 work.
- **Channeleon:** Developed a novel Deep Neural Network compression mechanism based on group channel on-line clustering that reduces the activation bit-widths to as low as 4 bits while outperforming the best state-of-the-art data-free methods' top-1 accuracy by 60%.

Data Engineering
- **CloudVision.ml:** Developed on the edge and in the cloud object detection and image classification with state of the art DNN models and cloud services. The project is used for waste management and traffic monitoring in urban area by analyzing outdoor live camera footages and won the 3rd place in Microsoft Encode competition.
- **Knowledge graph:** Designed and Implemented a software system for classifying the research papers stored in the cloud and build a knowledge graph of arXiv papers using a graph database. This framework helped ML researchers to execute complex queries revealing relation among publication and authors and conduct background reviews in the rapidly growing field.

Approximate Computing
- **Error Estimation:** Developed and implemented an accurate method for estimating errors in approximate circuits. Inputs are modelled as Gaussian mixtures and analytically propagated through circuit elements. Error estimates (MSE) are 2 orders of magnitude closer to actual simulations compared to prior approaches. Work published in *DATE 2017*.
- **Automatic Approximation:** Developed a software framework to automatically replace functional modules in Verilog RTL descriptions with their corresponding approximate modules and evaluate the overall design impact. The tool leverages Yosys, VTR, and Design Compiler.

A W A R D S

[2015] UBC Graduate Student Tuition Awards.　　　　　*[2011]* Awarded talented students' admission to M.Sc. program, IUST.

[2011] Ranked **2nd** in computer engineering dept. of IUST.　　*[2010]* Distinguished student of the computer engineering dept., IUST.

[2006] Top-10 in Sharif University's robotics competitions.

T E A C H I N G

Teaching Assistant:

[2015-2020] Capstone Design Projects (4×)　　*[2017-2019]* Digital Systems Design (4×)　　*[2018-2019]* Intro. to Digital Systems (2×)

[2019] Electrical Eng. Design Studio　　*[2015-2016]* Intro. to Microcomputers (2×)　　*[2013]* Advanced Computer Architecture

Course Development Grant:

[2019] **TLEF**: Teaching and Learning Enhancement Fund Project

***Certificates and Workshops*:**

[2019] CIRTL Associate, UBC.　　　　*[2019]* Foundations of Pedagogy　　　　*[2018]* Instructional Skill Workshops (ISW)

P U B L I C A T I O N S

- D. Yang, **A. Ghasemazar**, X. Ren, M. Golub, G. Lemieux, M. Lis, "Procrustes: A Dataflow and Accelerator for Sparse Training", In International Symposium on Microarchitecture (*MICRO*), 2020.
- **A. Ghasemazar**, P. Nair, M. Lis, "Thesaurus: Efficient Cache Compression via Dynamic Clustering," In The Int'l Conference on Architectural Support for Programming Languages and Operating Systems (*ASPLOS*), 2020.
- **A. Ghasemazar**, M. Ewais, P. Nair, M. Lis, "2DCC: Cache Compression in Two Dimensions," In Design, Automation and Test in Europe Conference (*DATE*), 2020.
- **A. Ghasemazar**, M. Lis, "Gaussian Mixture Error Estimation for Approximate Circuits," In Design, Automation and Test in Europe Conference and Exhibition (*DATE*), 2017.
- M. Kamal, **A. Ghasemazar**, A. Afzali-Kusha, M. Pedram, "Improving efficiency of extensible processors by using approximate custom instructions," Design, Automation and Test in Europe Conference and Exhibition (*DATE*), 2014.

R E F E R E N C E S

Dr. Mieszko Lis
Assistant Professor - *School of Electrical and Computer Engineering*
The University of British Columbia
Phone: +1(604)827-0738 | e-mail: mieszko@ece.ubc.ca

Dr. Prashant Nair
Assistant Professor - *School of Electrical and Computer Engineering*
The University of British Columbia
Phone: +1(604)827-0079 | e-mail: prashantnair@ece.ubc.ca