William S. Moses

wsmoses.com | wsmoses@illinois.edu | +1 (703) 638-2387

Research Vision: My research focuses on building compilers that automatically enables researchers to leverage the latest developments in high-performance computing and machine learning.

Education

Massachusetts Institute of Technology (MIT)

Ph.D., Computer Science, *Advisor: Charles E. Leiserson* — Aug 2017 – Jun 2023
Thesis: Compiler Abstractions and Transformations to Reduce Programming Burden

M.Eng., EECS, *Advisor: Charles E. Leiserson* — Jan 2017 – Jun 2017
Thesis: How Should Compilers Represent Fork-Join Parallelism?

B.S., EECS & Physics — Aug 2014 – Jun 2017

Research Appointments

University of Illinois, Urbana-Champaign, Urbana, IL

Assistant Professor, Computer Science — Aug 2024 – present

Assistant Professor, Electrical and Computer Engineering (courtesy) — Aug 2024 – present

Assistant Professor, Coordinated Science Laboratory — Aug 2024 – present

Adjunct Professor, Computer Science — Aug 2023 – present

Google Deepmind, Cambridge, MA — Oct 2023 – present
Visiting Researcher
Conduct research on compiler abstractions for efficient machine learning

University of Texas, Austin, Austin, TX — Sep 2023 – Dec 2023
J. Tinsley Oden Faculty Fellow
Conduct research on applying high performance computing and machine learning to problems in climate science.

MIT Computer Science and Artificial Intelligence Lab, Cambridge, MA — Sep 2014 – present
Researcher, Supercomputing Technologies Group
Compilers for parallelism (Tapir/Polygeist), differentiation (Enzyme), ML (TC), polyhedral model (TC/Polygeist), encryption (Syfer), phase-ordering (AutoPhase/ProTuner), scheduling (AutoPhase/ProTuner), cloud-compilation (Cymbl) & more.

Argonne National Laboratory, (virtual) Chicago, IL — Jan 2021 – Jul 2021
Researcher
Built parallel-specific optimizations and extended Enzyme to automatically synthesize gradients for CUDA, AMD, OpenMP, and MPI programs. Papers at SC '21, SC '22 and in submission to PLDI '23.

Lawrence Berkeley National Laboratory, Berkeley, CA — May 2019 – Aug 2019
Researcher, Quantum Algorithms Group
Created probabilistic programming framework for modeling quantum circuits. Posters at IWQC'19 & 2020 APS Meeting.

Facebook AI Research Laboratory (FAIR), New York, NY; Paris, France — Jun 2017 – Aug 2017
Research Intern
Performance engineering of deep neural networks by creating tensor DSL for kernel fusion and specialization for use in PyTorch and similar frameworks. Paper in ACM TACO journal.

Space Exploration Technologies (SpaceX), Hawthorne, CA — Jun 2015 – May 2016
Software / Propulsion Engineering Intern
Developed advanced distributed GPU multiphysics simulation using Wavelet Compression. Paper at SC'15 visualization.

U.S. Naval Research Laboratory, Washington, DC — Jun 2013 – Aug 2014
Electrical Engineering Intern
Created machine-learning algorithm to identify gaps in wireless spectrum to improve signal bandwidth and resilience.

Awards & Grants

Artificial Intelligence Accelerator Transition Award, US Air Force 2023
for the successful design implementation, and community building around he Enzyme automatic differentiation project

Julia Community Prize 2023
for significant contribution to the Julia language, community, and/or ecosystem for the development of Enzyme.jl

1st place, Student Research Competition at CGO 2023
for my poster/presentation/paper on "HTO: Header Time Optimization".

Great Dome Award 2023
for my event work as part of the 2018 MegaPi committee to organize the largest and most successful Pi reunion fundraising event, with over 2000 attendees, and $350,000 budget.

DOE Extreme-Scale Science Grant, *Differentiating Large-Scale Finite Element Applications* 2022
Co-investigator of a 2-year $900,000 grant from the Department of Energy to combine high-level information from finite element method applications within my Enzyme automatic differentiation tool.

Best Student Paper Award at SC 2022
Best Paper Finalist at SC (top 7 of 81 accepted and 320 submitted) 2022
for paper "Scalable Automatic Differentiation of Multiple Parallel Paradigms through Compiler Augmentation"

Karl Taylor Compton Prize, MIT's highest student award 2021
for my work to foster a safe and inclusive environment for computing, enable privacy-conscious remote-learning, build pandemic infrastructure and entertainment (featured in Business Insider & MIT Museum), and more.

Golden Beaver Award 2021
for creating positive change within the MIT computing student group, SIPB.

Best Student Paper Finalist at SC (top 5 of 98 accepted and 365 submitted) 2021
Best Reproducability Advancement Finalist at SC (top 5 of 98 accepted and 365 submitted) 2021
for paper "Reverse-Mode Automatic Differentiation and Optimization of GPU Kernels via Enzyme"

Spotlight Paper (top 15% accepted) at NeurIPS 2020
Best Student Talk at US LLVM Dev Meeting 2020
for paper "Instead of Rewriting Foreign Code for Machine Learning, Automatically Synthesize Fast Gradients"

Best Student Talk (tie) at US LLVM Dev Meeting 2019
for talk "HTO: Header-Time Optimization"

Department of Energy Computational Science Fellowship 2018-2022
fully funded PhD stipend, tuition & benefits for 4 years (∼$450K)

National Science Foundation Graduate Research Fellowship, declined 2018

Best Paper Award at PPoPP 2017
for paper "Tapir: Embedding Recursive Fork-Join Parallelism into LLVM's Intermediate Representation"

Robert M. Fano Award 2017
for outstanding capstone undergraduate research project, Tapir compiler for parallel programs.

Larry G. Benedict Award 2017
MIT-wide award for mentorship and "empowering fellow students to develop as leaders"

Finalist, MIT Idea2 Global Competition 2016
one of 15 finalists in a worldwide biotech idea competition sponsored by MIT linQ, AstraZeneca, Fipse, and PDS

1st place, Innovative Defensive Technologies Programming Contest 2014
month-long competition to develop software capable of automatically identifying errors in mission-critical systems

Gold (highest) Division, USA Computing Olympiad 2012-2014
premier high school programming competition; Gold was the highest division while I was a competitor

Selected Publications

Over 20 publications in various journals (TOPC, TACO), book chapters, conferences (SC, PACT, NeurIPS, MLSys, PPoPP), and workshops (IMPACT, NeurIPS Robust AI in FS, PPoPP-PMAM, SCVis, LLVMHPC, IEEE FCCM). Full list at end and on website (`https://wsmoses.com`).

SC '22 — *Scalable Automatic Differentiation of Multiple Parallel Paradigms through Compiler Augmentation*
William S. Moses, Sri Hari Krishna Narayanan, Ludger Paehler, Valentin Churavy, Jan Hückelheim, Michel Schanen, Johannes Doerfert, and Paul Hovland
Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, **Best Student Paper Award & Best Paper Finalist**

PACT '21 — *Polygeist: Raising C to Polyhedral MLIR*
William S. Moses, Lorenzo Chelini, Ruizhe Zhao and Oleksandr Zinenko
Proceedings of the ACM International Conference on Parallel Architectures and Compilation Techniques

SC '21 — *Reverse-Mode Automatic Differentiation and Optimization of GPU Kernels via Enzyme*
William S. Moses, Valentin Churavy, Ludger Paehler, Jan Hückelheim, Sri Hari Krishna Narayanan, Michel Schanen and Johannes Doerfert
Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, **Best Student Paper Finalist**

NeurIPS '20 — *Instead of Rewriting Foreign Code for Machine Learning, Automatically Synthesize Fast Gradients*
William S. Moses, Valentin Churavy
Advances in Neural Information Processing Systems Vol 33, **Spotlight Paper (top 15% accepted)**

TACO '19 — *The Next 700 Accelerated Layers: From Mathematical Expressions of Network Computation Graphs to Accelerated GPU Kernels, Automatically (Tensor Comprehensions)*
Nicolas Vasilache, Oleksandr Zinenko, Theodoros Theodoridis, Priya Goyal, Zachary DeVito, **William S. Moses**, Sven Verdoolaege, Andrew Adams, and Albert Cohen
Journal, ACM Transactions on Architecture and Code Optimization

PPoPP '17 — Tapir: Embedding Fork-Join Parallelism in LLVM's Intermediate Representation
Tao B. Schardl, **William S. Moses**, Charles E. Leiserson
Proceedings of the 22nd ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming, **Best Paper Award**

Mentorship Experience

Master's Projects: *Tim Gymnich* (Technical University of Munich), compiler-based forward mode automatic differentiation, co-supervised with Ludger Paehler

MIT Undergraduate Research Program: *Carl Guo*, database compilation; *Sage Simhon*, quantum compiling & provably correct ML; *Douglas Kogut, Jiahao Li, Bojan Serafimov*, parallel optimizations and program representations

MIT PRIMES (high-school research program): *Carl Guo*, machine learning for automatic program optimization; *Sanath Govindarajan* (now undergraduate at UT Austin), synthesis of fast homomorphic encryption via MLIR; *Walden Yan* (now undergraduate at Harvard), Neural Program Synthesis

Google Summer of Code: *Chuyang Chen* alias analysis via lowered Rust types; *Pratush Das* (2020, now PhD student at Purdue), BLAS automatic differentiation, *Manuel Drehwald* (undergraduate at Karlsruhe Institute of Technology), automatic differentiation in Rust; *Shakil Ahmed*, new pass manager; *Abhishek Vu*, parallel optimizations.

Teaching Experience

Teaching Assistant, Introduction to Algorithms, MIT EECS — Feb 2018 – May 2018

Led and prepared recitations for MIT's introduction to algorithms, and wrote problem sets and exams. 350 total students, 25 in my section. Wrote problem sets and exams for all students. Rated 6.5/7.0 on student evaluations.

Lead Instructor, Introduction to C/C++, MIT EECS — Jan – Feb 2015, 2016, 2021

With one other person, created introductory undergraduate course on C/C++ (6.179). Presented and created lectures, weekly homework, and a final project. Over 200 students from MIT, Harvard, and Wellesley.

Conference Service

Co-Organizer, *NeurIPS Differentiable Programming Workshop* Aug 2021 – Dec 2022

Co-lead organizer of a new workshop at NeurIPS 2021. Ran logistics including: call for papers, organizing a program committee (reviewers), drafting a code of conduct, managing corporate sponsorship, and creating initiatives for DEI.

Co-Organizer, *Workshop on forward methods at SIAM CSE 2023* Aug 2022-present

Co-Organizer, *Mini-Workshop: Differentiable Programming for High-Performance, Data-Intensive Computations* Apr 2021

External Reviewer

2023 Conference on Machine Learning and Systems (MLSys), 2021 ACM SIGGRAPH, 2021 ACM SPAA

Program Committee Member,

1st Workshop on AI Assisted Software Development for HPC (AI4DEV at SC '24), 28th Workshop on High-Level Parallel Programming Models (HIPS at IPDPS '23), 2022 Workshop on LLVM in Parallel Processing at ICPP, 2021 NeurIPS ML For Systems Workshop, 2021 Workshop on LLVM in Parallel Processing at ICPP, 2020 US LLVM Developer's Meeting

Other Service

C++ Standards Committe; WG19 on Machine Learning Nov 2020 – present

Review proposals for additions to the C++ standard that relate to machine learning. Proposed language-level integration for automatic differentiation in technical paper.

Rust Machine Learning Working Group Nov 2020 – present

Review machine-learning proposals to the Rust language. Work on automatic differentiation and efficient neural-networks.

Reviewer, *MIT Engineering Advisory Board (EAB)* Oct 2020 – present

Read and review technical portfolios of prospective MIT undergraduate students.

The Engine Startup Accelerator Working Group Dec 2016 – Apr 2018

Helped found MIT's Engine hard-tech startup incubator. Worked with MIT's Technology Licensing Office, professors, start ups, and venture capitalists to recommend framework for intellectual property.

University Leadership

Alumni Class President, *MIT Class of 2018* June 2023 – present

President of the MIT class of 2018, representing the the class to MIT and Alumni Association, plan reunions, fundraising events, talk seminars, and work with admissions.

Chair, *Student Information Processing Board (SIPB)* Feb 2020 – Feb 2021

Chair of MIT's oldest computer science club. Provide students and the community access to computing infrastructure (virtual machines, web hosting, chat bots), teaching infrastructure (a lecture series), and advocacy.

Graduate Community Fellow, *Institute Community and Equity Office* Sep 2017 – Jan 2018

Promoted diversity and inclusion at MIT through speaker series, events, news articles, and more.

Elected Councillor, *MIT Undergraduate Association* Feb 2016 – Jan 2017

One of twenty voting members of MIT's institute-wide student government. Working on issues such as the creation of a new dormitory, institute-wide sustainability programs, student withdrawal policy, among others.

Executive Board, *MIT Society of Physics Students* May 2016 – May 2017

Member of the executive board for the MIT chapter of SPS, the premier physics club for college students. Help run faculty dinners and organizes events such as the MIT undergraduate physics conference.

Executive Board, *MIT IEEE/ACM* May 2015 – May 2016

Member of the executive board for the MIT chapter of IEEE/ACM, the premier electrical engineering and computer science club for college students. Helped lead the first ever IEEE-sponsored conference for undergraduate students.

Research Projects

Enzyme, Cross-platform efficient compiler automatic differentiation — Jan 2019 – present

High-performance compiler for producing derivatives of arbitrary code, including GPU, OpenMP, and MPI. Achieves state of the art runtime and demonstrates 4.2×speedup on standard benchmarks, 477× on Swift, and two orders of magnitude on Julia . Integrations in PyTorch and TensorFlow enable automatic importing of external libraries. In use by Facebook, Julia Computing, Argonne/Lawrence Livermore National Labs, NASA, PassiveLogic, & others. More information at enzyme.mit.edu. Components published in NeurIPS 2020, SC 2021, and SC 2022.

Polygeist, Polyhedral Compiler & C/C++ frontend for MLIR — Oct 2019 – present

Novel C/C++ compiler for MLIR that also provides polyhedral transformations within MLIR. Novel optimizations such as statement splitting enables Polygeist to outperform existing tools and automatically parallelize programs. Ongoing work to transcompile existing GPU programs for other architectures, including CPU's. Collaboration with Riken to run GPU codes on Fugaku CPU-only supercomputer. In use by EPFL, ETH Zurich, Imperial College London, Intel, Google, SiFive, and Xilinx. See polygeist.mit.edu for more information. Components published in PACT 2021 and PPoPP 2023.

Tapir/LLVM, Parallel Compiler IR — Jan 2015 – present

Modification to LLVM IR in order to represent task-based parallelism. Performs both serial and parallel-specific optimizations on parallel tasks, resulting in significant performance boosts. See github.com/wsmoses/Tapir-LLVM for more information. Integrated into OpenCilk, LANL's Kitsune compiler, and used in MIT courses. Best paper at PPoPP 2017.

Tensor Comprehensions, Deep Learning Compilation Flow — May 2017 – Feb 2018

End-to-end compilation flow that live generates optimal GPU kernels for machine learning frameworks by leveraging the polyhedral compilation framework. See facebookresearch.github.io/TensorComprehensions for more information. Published in TACO.

Auto-Phase/Pro-Tuner, Smart Compiler — Sep 2018 – 2020

Reinforcement learning tools for deriving optimal optimization pass orderings / schedules. Published in MLSys 2020.

Header Time Optimization (HTO), Whole program optimization framework — Jun 2019 – present

Whole-codebase optimization tool achieving 50% of the speedups of existing tools without any additional cost. See github.com/wsmoses/LLVM-HTO for more information. Best student presentation at 2019 LLVM Dev Meeting.

Cymbl, Cloud compiler — Nov 2019 – present

Infinitely scalable caching cloud compiler. State-of-the art compile times and up to 500×faster than serial compilation. See cymbl.dev for more information.

Quantum Bayes, Bayesian Framework for Learning Quantum Error Models — Jun 2019 – present

Extended quantum simulators to be compatible with existing probabilistic programming frameworks to create a system capable of learning the most probable quantum error model given experimental data. Presented posters at IWQC'19 and 2020 APS March Meeting.

SyFER, Fast Homomorphic Encrpyion Library — Jan 2018 – present

Fast library for fully homomorphic encryption, with an emphasis on performing secure machine learning.

LiTM, LightWeight Transactional Memory — Aug 2017 – July 2019

Tool for automatically parallelizing loops through the use of transactional memory. Published in PMAM 2019.

Flamingo, Multi-resolution physics simulator — Jun 2015 – May 2016

A multi-resolution physics simulator based on wavelets capable of accurately predicting the behavior of combustible fluids. In use by SpaceX for development of rocket engines. Published at SC Visualization Challenge 2015.

Journal Papers (Peer-Reviewed)

[18] T. B. Schardl, **Moses, William S.**, and C. E. Leiserson, "Tapir: Embedding recursive fork-join parallelism into LLVM's intermediate representation," *ACM Trans. Parallel Comput.*, vol. 6, no. 4, Dec. 2019, ISSN: 2329-4949. DOI: 10.1145/3365655.

[20] N. Vasilache, O. Zinenko, T. Theodoridis, P. Goyal, Z. Devito, **Moses, William S.**, S. Verdoolaege, A. Adams, and A. Cohen, "The next 700 accelerated layers: From mathematical expressions of network computation graphs to accelerated GPU kernels, automatically," *ACM Trans. Archit. Code Optim.*, vol. 16, no. 4, Oct. 2019. DOI: 10.1145/3355606.

Conference Papers (Peer-Reviewed)

[5] **Moses, William S**, I. R. Ivanov, J. Domke, T. Endo, J. Doerfert, and O. Zinenko, "High-performance GPU-to-CPU transpilation and optimization via high-level parallel constructs," in *Proceedings of the 28th ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming*, Montreal, Canada: ACM, 2023. [Online]. Available: https://arxiv.org/pdf/2207.00257.pdf.

[6] M. Schanen, S. H. K. Narayanan, S. Williamson, V. Churavy, **Moses, William S.**, and L. Paehler, "Transparent checkpointing for automatic differentiation of program loops through expression transformations," in *Computational Science – ICCS 2023*, J. Mikyška, C. de Mulatier, M. Paszynski, V. V. Krzhizhanovskaya, J. J. Dongarra, and P. M. Sloot, Eds., Cham: Springer Nature Switzerland, 2023, pp. 483–497, ISBN: 978-3-031-36024-4.

[8] Z. C. Guo and **Moses, William S.**, "Enabling transformers to understand low-level programs," in *2022 IEEE High Performance Extreme Computing Conference (HPEC)*, 2022, pp. 1–9. DOI: 10.1109/HPEC55821.2022.9926313.

[9] **Moses, William S**, S. Hari Krishna Narayanan, L. Paehler, V. Churavy, J. Hückelheim, M. Schanen, J. Doerfert, and P. Hovland, "Scalable automatic differentiation of multiple parallel paradigms through compiler augmentation," in *SC '22: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, ACM, 2022. [Online]. Available: https://c.wsmoses.com/papers/enzymePar.pdf. ***Best Student Paper Award and Best Paper Finalist***

[10] **Moses, William S**, V. Churavy, L. Paehler, J. Hückelheim, S. Hari Krishna Narayanan, M. Schanen, and J. Doerfert, "Reverse-mode automatic differentiation and optimization of GPU kernels via enzyme," in *SC '21: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, ACM, 2021. DOI: 10.1145/3458817.3476165. ***Best Student Paper Finalist and Best Reproducibility Advancement Finalist***

[12] **Moses, William S.**, L. Chelini, R. Zhao, and O. Zinenko, "Polygeist: Raising C to polyhedral MLIR," in *Proceedings of the ACM International Conference on Parallel Architectures and Compilation Techniques*, Virtual Event: Association for Computing Machinery, 2021. [Online]. Available: https://c.wsmoses.com/papers/Polygeist_PACT.pdf.

[15] A. Haj-Ali, Q. J. Huang, J. Xiang, **Moses, William**, K. Asanovic, J. Wawrzynek, and I. Stoica, "Autophase: Juggling HLS phase orderings in random forests with deep reinforcement learning," *Proceedings of Machine Learning and Systems*, vol. 2, pp. 70–81, 2020. [Online]. Available: https://proceedings.mlsys.org/paper/2020/file/4e732ced3463d06de0ca9a15b6153677-Paper.pdf.

[16] **Moses, William** and V. Churavy, "Instead of rewriting foreign code for machine learning, automatically synthesize fast gradients," in *Advances in Neural Information Processing Systems*, 2020. [Online]. Available: https://dl.acm.org/doi/pdf/10.5555/3495724.3496770. ***Spotlight Presentation***

[23] Schardl, Tao B., **Moses, William S.**, and Leiserson, Charles E., "Tapir: Embedding fork-join parallelism into LLVM's intermediate representation," in *Proceedings of the 22nd ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming*, Austin, Texas, USA: ACM, Jan. 2017, pp. 249–265, ISBN: 978-1-4503-4493-7. DOI: 10.1145/3018743.3018758. ***Best Paper Award***

Workshop Papers (Peer-Reviewed)

[11] **Moses, William S.**, L. Chelini, R. Zhao, and O. Zinenko, "Polygeist: Affine C in MLIR," in *IMPACT 2021-11th International Workshop on Polyhedral Compilation Techniques*, 2021. [Online]. Available: https://acohen.gitlabpages.inria.fr/impact/impact2021/papers/IMPACT_2021_paper_1.pdf.

[17] Q. Huang, A. Haj-Ali, **Moses, William**, J. Xiang, I. Stoica, K. Asanovic, and J. Wawrzynek, "Autophase: Compiler phase-ordering for HLS with deep reinforcement learning," in *2019 IEEE 27th Annual International Symposium on Field-Programmable Custom Computing Machines (FCCM)*, IEEE, 2019, pp. 308–308. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8735549.

[19] Y. Shavit and **Moses, William S.**, "Extracting incentives from black-box decisions," in *2019 NeurIPS Workshop on AI in Financial Services*, 2019. [Online]. Available: https://arxiv.org/pdf/1910.05664.pdf.

[21] Y. Xia, X. Yu, **Moses, William**, J. Shun, and S. Devadas, "LiTM: A lightweight deterministic software transactional memory system," in *Proceedings of the 10th International Workshop on Programming Models and Applications for Multicores and Manycores*, ACM, 2019, pp. 1–10. [Online]. Available: https://c.wsmoses.com/papers/litm.pdf.

[24] G. Stelle, **Moses, William S.**, S. L. Olivier, and P. McCormick, "OpenMPIR: Implementing openmp tasks with tapir," in *Proceedings of the Fourth Workshop on the LLVM Compiler Infrastructure in HPC*, Denver, CO, USA: ACM, 2017, 3:1–3:12, ISBN: 978-1-4503-5565-0. DOI: 10.1145/3148173.3148186. [Online]. Available: https://c.wsmoses.com/papers/openmpir.pdf.

[25] **Moses, William S.**, T. B. Schardl, and C. E. Leiserson, "Embedding fork-join parallelism into llvm ir," in *19th Workshop on Compilers for Parallel Computing*, 2016. [Online]. Available: https://cpc2016.infor.uva.es/wp-content/uploads/2016/06/CPC2016_paper_12.pdf.

[26] J. Balme, E. Brown-Dymkoski, V. Guerrero, S. Jones, A. Kessler, A. Lichtl, K. Lung, **Moses, William**, K. Museth, N. Roberson, *et al.*, "Extreme multi-resolution visualization: A challenge on many levels," in *SuperComputing Visualization Contest 2015*, 2015. [Online]. Available: https://c.wsmoses.com/papers/spacex15.pdf.

Book Chapters (Peer-Reviewed)

[27] E. D. Demaine and **William S. Moses**, "Computational complexity of arranging music," in *Revised Papers from MOVES 2015: Mathematics of Various Entertaining Subjects*, Princeton University Press, 2015. [Online]. Available: https://c.wsmoses.com/papers/moves15.pdf.

Preprints

[4] J. Hückelheim, H. Menon, **Moses, William**, B. Christianson, P. Hovland, and L. Hascoët, "Understanding automatic differentiation pitfalls," *arXiv preprint arXiv:2305.07546*, arXiv, 2023.

[7] J. Brown, V. Barra, N. Beams, L. Ghaffari, M. Knepley, **Moses, William**, R. Shakeri, K. Stengel, J. L. Thompson, and J. Zhang, "Performance portable solid mechanics via matrix-free p-multigrid," *arXiv preprint arXiv:2204.01722*, 2022. [Online]. Available: https://arxiv.org/pdf/2204.01722.pdf.

[13] S. Govindarajan and **Moses, William S**, *SyFER-MLIR: Integrating fully homomorphic encryption into the MLIR compiler framework*, 2020. [Online]. Available: https://math.mit.edu/research/highschool/primes/materials/2020/Govindarajan-Moses.pdf.

[14] A. Haj-Ali, H. Genc, Q. Huang, **Moses, William**, J. Wawrzynek, K. Asanovi⊠, and I. Stoica, "Protuner: Tuning programs with monte carlo tree search," *arXiv preprint arXiv:2005.13685*, 2020. [Online]. Available: https://arxiv.org/pdf/2005.13685.pdf.

[22] N. Vasilache, O. Zinenko, T. Theodoridis, P. Goyal, Z. DeVito, **Moses, William S**, S. Verdoolaege, A. Adams, and A. Cohen, "Tensor comprehensions: Framework-agnostic high-performance machine learning abstractions," *arXiv preprint arXiv:1802.04730*, 2018. [Online]. Available: https://arxiv.org/pdf/1802.04730.pdf.

Invited Talks

Colloquium Talks

MIT Thesis Defense — May 2023
'Supercharging Programming Through Compiler Technology'

Mathworks Code Generation Seminar — Jan 2023
High-Performance GPU-to-CPU Transpilation and Optimization

BU Systems Seminar — Dec 2022
Enzyme: High-Performance, Cross-Language, and Parallel Automatic Differentiation

Columbia DSI Seminar — Jun 2022
Enzyme: High-Performance, Cross-Language, and Parallel Automatic Differentiation

ExaSGD Seminar — Jun 2022
Updates on Enzyme: High-Performance, Cross-Language, and Parallel Automatic Differentiation

TUM Seminar — Jun 2022
Enzyme: High-Performance, Cross-Language, and Parallel Automatic Differentiation

Google/INRIA/ONERA AD Meeting — May 2022
Enzyme: High-Performance, Cross-Language, and Parallel Automatic Differentiation

Imperial College London Seminar — May 2022
Enzyme: High-Performance, Cross-Language, and Parallel Automatic Differentiation

NVIDIA Seminar — Feb 2022
Enzyme: High-Performance, Cross-Language, and Parallel Automatic Differentiation

LLNL Invited Seminar — Dec 2021
Enzyme: High-Performance Automatic Differentiation of LLVM

Washington University of St. Louis Colloquium — Nov 2021
Enzyme: High-Performance, Cross-Language, and Parallel Automatic Differentiation

CU Boulder CS Colloquium — Oct 2021
Enzyme: High-Performance, Cross-Language, and Parallel Automatic Differentiation

Tobias Grosser Group Meeting (Edinburgh) — Aug 2021
Polygeist: Raising C to Polyhedral MLIR

Legion Group Meeting (Stanford) — Jun 2021
Instead of Rewriting Foreign Code for Machine Learning, Automatically Synthesize Fast Gradients!

Jiantao Jiao Group Meeting (Berkeley) — Jun 2021
Instead of Rewriting Foreign Code for Machine Learning, Automatically Synthesize Fast Gradients!

CaaS Monthly Meeting (Princeton) — Jun 2021
Cymbl: To -jInfinity & Beyond

Apple Ted-K Talk — Nov 2020
Cymbl: To -jInfinity & Beyond

Secure AI Labs Seminar Series Jul 2020

Making ML Fast for Arbitrary Code (Enzyme)

Argonne National Laboratories Seminar Jul 2020

Post-Optimization Automatic Differentiation by Synthesizing LLVM

Conference Talks

CGO SRC 2023 , **ACM Gold Award (1st place)** Feb 2023

'HTO: "Header"-Time Optimization'

PPoPP 2023 Feb 2023

High-Performance GPU-to-CPU Transpilation and Optimization via High-Level Parallel Constructs

SC 2022 , **Best Student Paper** Nov 2022

Scalable Automatic Differentiation of Multiple Parallel Paradigms through Compiler Augmentation

SC '21 (The International Conference for High Performance Computing, Networking, Storage, and Analysis) , **Best Student Paper Finalist and Best Reproducibility Advancement Finalist** Nov 2021

Reverse-Mode Automatic Differentiation and Optimization of GPU Kernels via Enzyme

CPPCon 2021 Oct 2021

Differentiable Programming in C++

PACT Conference 2021 Sep 2021

Polygeist: Raising C to Polyhedral MLIR

NVIDIA GTC 2021 Apr 2021

Post-Optimization Automatic Differentiation by Synthesizing LLVM

NeurIPS 2020, Spotlight Talk , **Spotlight** Dec 2020

Instead of Rewriting Foreign Code for Machine Learning, Automatically Synthesize Fast Gradients

Rework Deep Learning Summit Boston 2018 May 2018

Tensor Comprehensions

Mathematics of Various of Entertaining Subjects (MOVES) 2015 Aug 2015

Computational Complexity of Arranging Music

Workshop Talks

AI4Dev @ SC'23 , **Keynote** Nov 2023
'LLVM in the age of LLMs: Machine Learning for IR and optimization and more'

MLIR Summit @ LLVM Dev Meeting Oct 2023
'Enzyme-MLIR: Early Experiments on multi-level automatic differentiation'

Differentiable and Probabalistic Programming for Fundamental Physics Jun 2023
'An Introduction to Enzyme and Some Fun Recent Results'

EuroAD 2023 Jun 2023
'Recent Compiler-Based AD Results and Open Questions'

Enzyme Conference 2023 Feb 2023
Enzyme Tutorial

LLVM HPC @ SC 2022 , **Keynote Talk** Nov 2022
Polygeist C++ frontend for MLIR

MLIR Summit @ 2022 US LLVM Dev Meeting Nov 2022
Polygeist C++ frontend for MLIR

Second MODE Workshop on Differentiable Programming for Experiment Design Sep 2022
Synthesization of Fast Gradients with Enzyme

LLPP '22 Aug 2022
Enzyme: Automatic Differentiation for Parallel Programs

JuliaCon ESM MiniSymposium Jul 2022
Enzyme.jl

RSS '22 Workshop on Differential Simulation Jul 2022
Automatic Differentiation of Black Box Code with Enzyme

ISC LLVM Performance Workshop Jun 2022
MLIR-In-The-Middle: compiling C++ and extensions via the new extensible infrastructure

LLVM Performance Workshop at CGO '22 Apr 2022
[Tutorial] An Guide to Performance Debugging LLVM-based Programs

SIAM PP22 GPU MiniSymposium Jan 2022
Reverse-Mode Automatic Differentiation and Optimization of GPU and Heterogeneous Parallel Programs via Enzyme

Virtual LLVM Developer Meeting, Fall 2021 Nov 2021
How to Use Enzyme to Automatically Differentiate Any LLVM-based Language for CPU, GPU, and More

7th Annual Workshop on the LLVM Compiler Infrastructure in HPC Nov 2021
Enzyme: Fast, Language Agnostic, Differentiation of Parallel Programs in LLVM

European Workshop on Automatic Differentiation 2021 Nov 2021
Language-Independent Automatic Differentiation and Optimization of GPU Programs with Enzyme

Differentiable Programming Workshop Apr 2021
Post-Optimization Automatic Differentiation by Synthesizing LLVM

IMPACT 2021 — Jan 2021
Polygeist: Affine C in MLIR

Languages For Inference (LAFI) 2021 — Jan 2021
Enzyme: High-Performance Automatic Differentiation of LLVM

US LLVM Developer Meeting, Fall 2020 , **Best Student Presentation** — Oct 2020
Enzyme: High-Performance Automatic Differentiation of LLVM

European Workshop on Automatic Differentiation 2020 — Aug 2020
Post-Optimization Automatic Differentiation by Synthesizing LLVM

Fourth LLVM Performance Workshop at CGO , **Keynote Talk** — Feb 2020
Header Time Optimization: Cross-Translation Unit Optimization via Annotated Headers

3rd International Workshop on Quantum Compilation — Nov 2019
Automated Bayesian Estimation of Quantum Error Models

US LLVM Developer Meeting, Fall 2019 , **Best Student Presentation (Tie)** — Oct 2019
"Header Time Optimization": Cross-Translation Unit Optimization via Annotated Headers

European Workshop on Automatic Differentiation 2019 — Jul 2019
Enzyme: Efficient Cross-Platform AD by Synthesizing LLVM

US LLVM Developer Meeting, Fall 2018 — Oct 2018
How to Use LLVM To Optimize Parallel Programs

LLVM Workshop at CGO 2018 — Feb 2018
Tensor Comprehensions

US LLVM Developer Meeting, Fall 2017 — Oct 2017
Leveraging LLVM to Optimize Parallel Programs

IBM PL Day 2016 — Dec 2017
Tapir: Embedding Fork-Join Parallelism into LLVM IR

Compilers for Parallel Computing 2016 — Jul 2016
Embedding Fork-Join Parallelism into LLVM IR

Other Talks

CESMIX TST '22 — May 2022
Enzyme: High-Performance, Cross-Language, and Parallel Automatic Differentiation

Cambridge Area Julia Users Network (CAJUN) — May 2022
A brief introduction to Enzyme.jl

MIT 18.065 Lecture — May 2022
Back Propagation and Automatic Differentiation

DJ4Earth — Mar 2022
Enzyme and Enzyme.jl Updates

MLIR Open Design Meeting — Feb 2021
Polygeist: Affine C in MLIR

MIT 6.S898 Lecture — Apr 2017
Tapir: Embedding Fork-Join Parallelism into LLVM IR

Intel Corporation — Jan 2015
Syntactic Simplifications for Reducer Hyperobjects