Artificial Intelligence and Machine Learning for High Energy Physics

Phiala Shanahan, MIT

Image Credit: 2018 EIC User's Group Meeting

2023 April P5 Town Hall, BNL

Massachusetts Institute of Technology

AI/ML impacts all facets of HEP

AI/ML is a class of computational tools with tremendous potential across HEP applications

- Experiment
- Data analysis
- Theory

i.e., HEP uses span beyond big-data applications

★ AI/ML is already integral in applications across HEP, many still in early stages.

Now is the crucial time to develop infrastructure and frameworks to enable maximal exploitation

AI/ML in HEP theory

Analytic

Formal theory
incl. string/gravity

First-principles theory
incl. lattice field theory

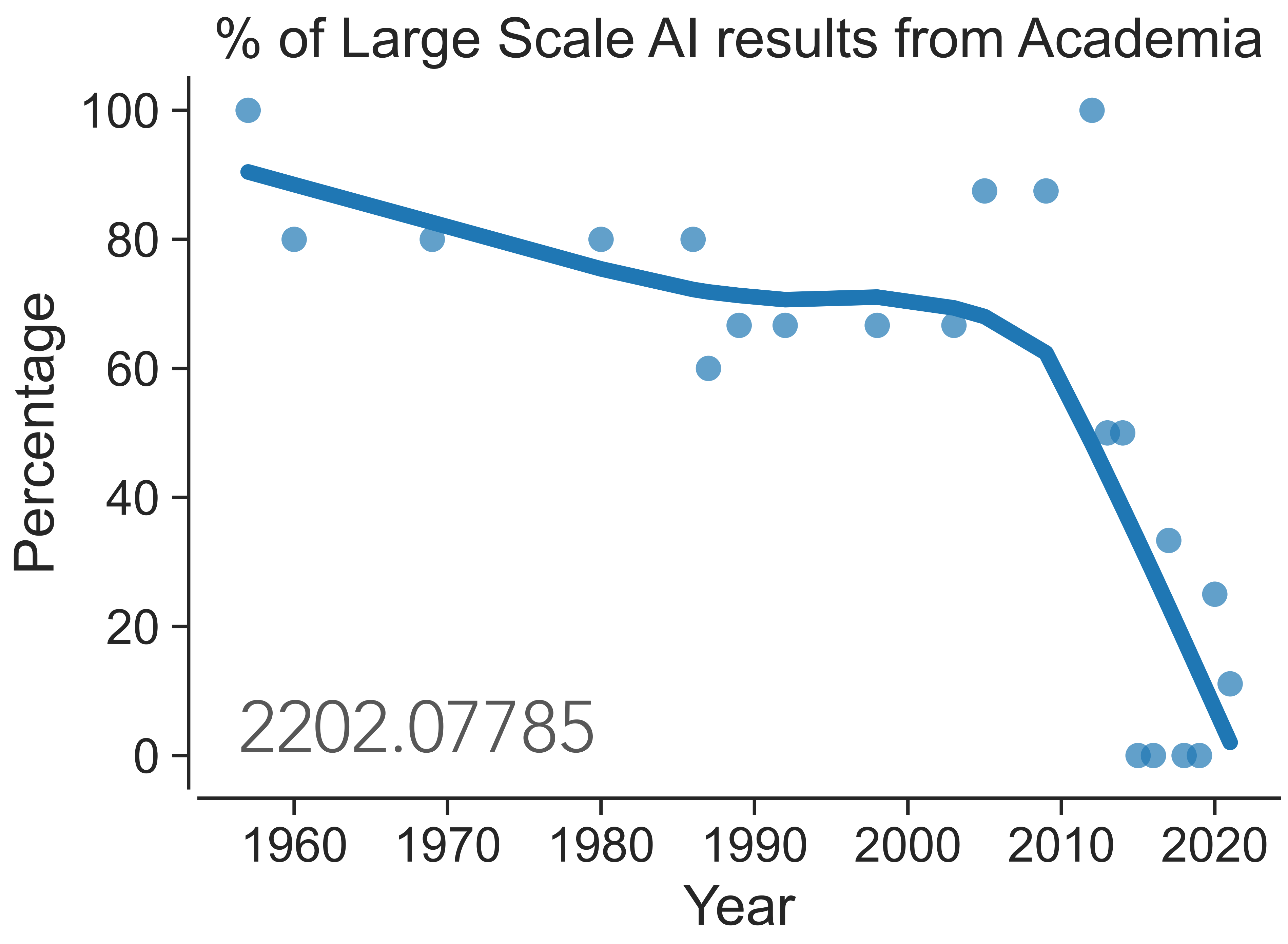
Particle pheno
and cosmology

Data analysis

Empirical

The landscape of AI/ML is rapidly changing

- **Industry now leads the way** in large AI models
- Large, general (foundation) models are not necessarily something we should (or can afford to) aspire to create as an academic community
- **Enabling scale is critical, but the race to exploit large models isn't the only frontier for HEP:** Applications are structured, with significant domain expertise/info to exploit, incl. symmetries, invariances, conservation laws, limits…



The development of HEP-specific AI/ML requires targeted investment and support

HEP as consumers and ***developers*** of AI/ML

Both exploitation and innovation in AI/ML will push HEP science forward

Exploitation

- Exploit general AI/ML developments at different levels
  - Build on large general models
  - Adapt AI/ML tools developed outside HEP to HEP problems
- Knowledge transfer into HEP

Innovation

- HEP-specific AI/ML designed for HEP-specific applications
  - Rapidly advancing as our community gains AI/ML literacy
  - Requires “bilingual” workforce
- Knowledge transfer out of HEP

HEP as consumers and ***developers*** of AI/ML

Long-term planning must cover extremes of scope and scale and adapt over time

- **Advances to be made at every level of complexity and scale**

  Complexity: Existing tools → Custom approaches

  Scale: Laptop → Exascale hardware

AURORA | EARLY SCIENCE PROGRAM FOR DATA & LEARNING

- Many applications are in an early phase of development
- We have not yet explored the full space of possibilities: **new paradigms certainly still to come in next decade**

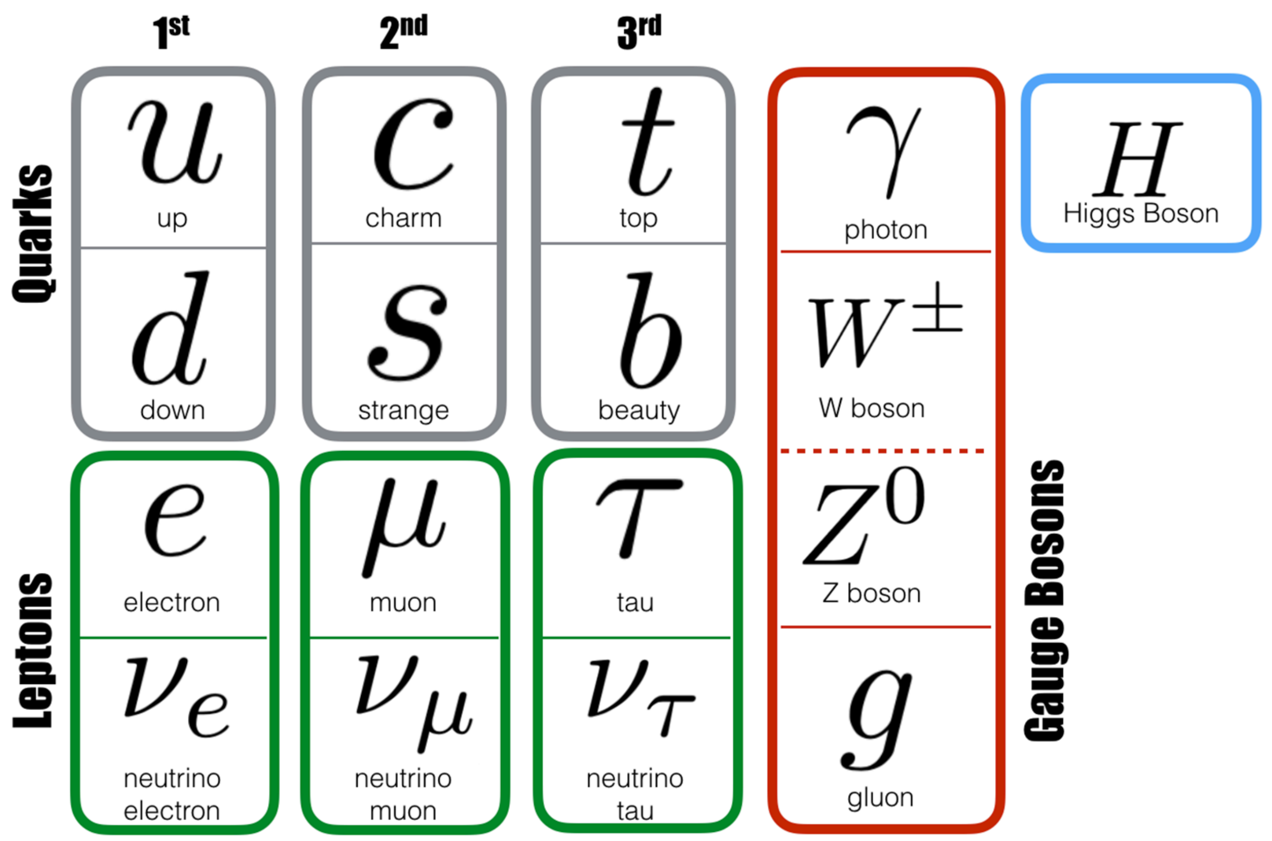
Applications across industry and HEP theory and experiment share some challenges but offer **unique demands and opportunities**

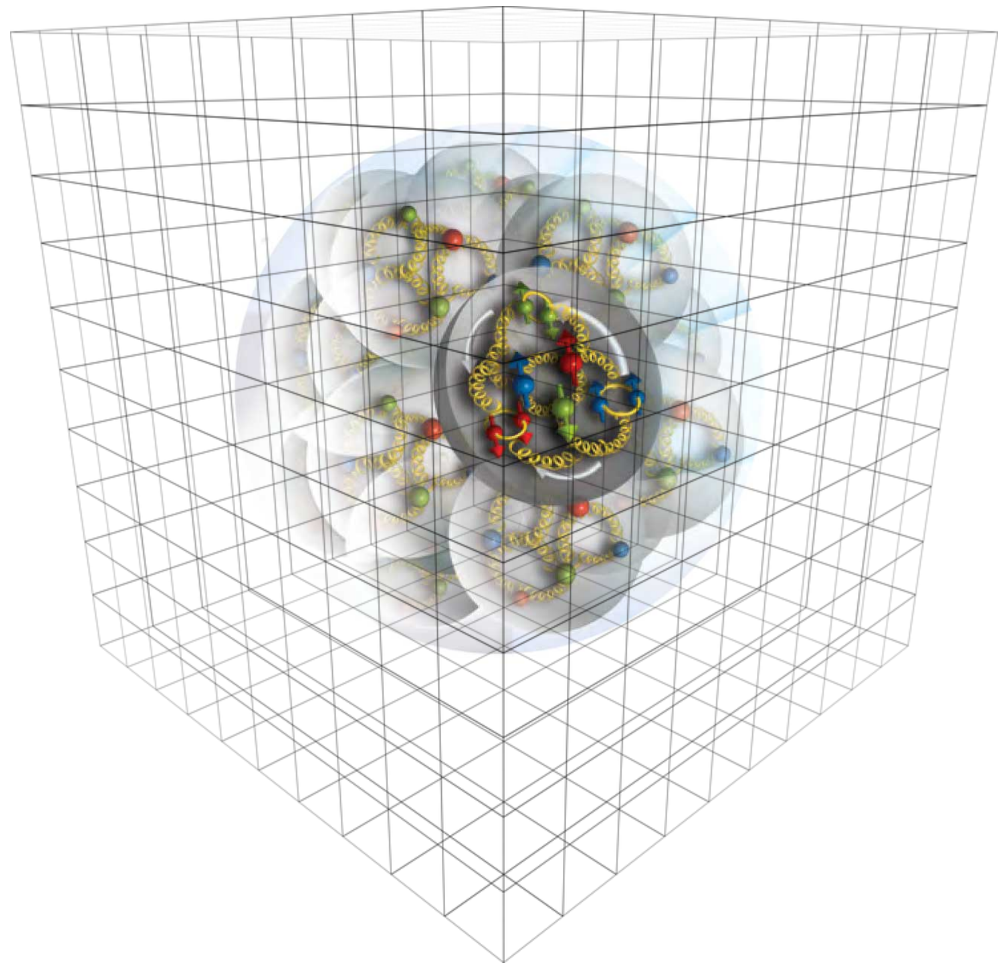
HEP-inspired AI/ML can have broad impact

Long history of HEP driving innovation leading to interdisciplinary advances!

Theory case study: **Lattice QCD**

Numerical first-principles approach to non-perturbative QCD calculations





Lattice QCD provides input for

- Decay constants, form factors, mixing parameters
- Hadronic vacuum polarisation and light-by-light scattering
- Neutrino-nucleus interactions
- Dark matter-nucleon and DM-nucleus interactions
- Muon-nucleus cross-sections
- Parton distribution functions

HEP-inspired AI/ML can have broad impact

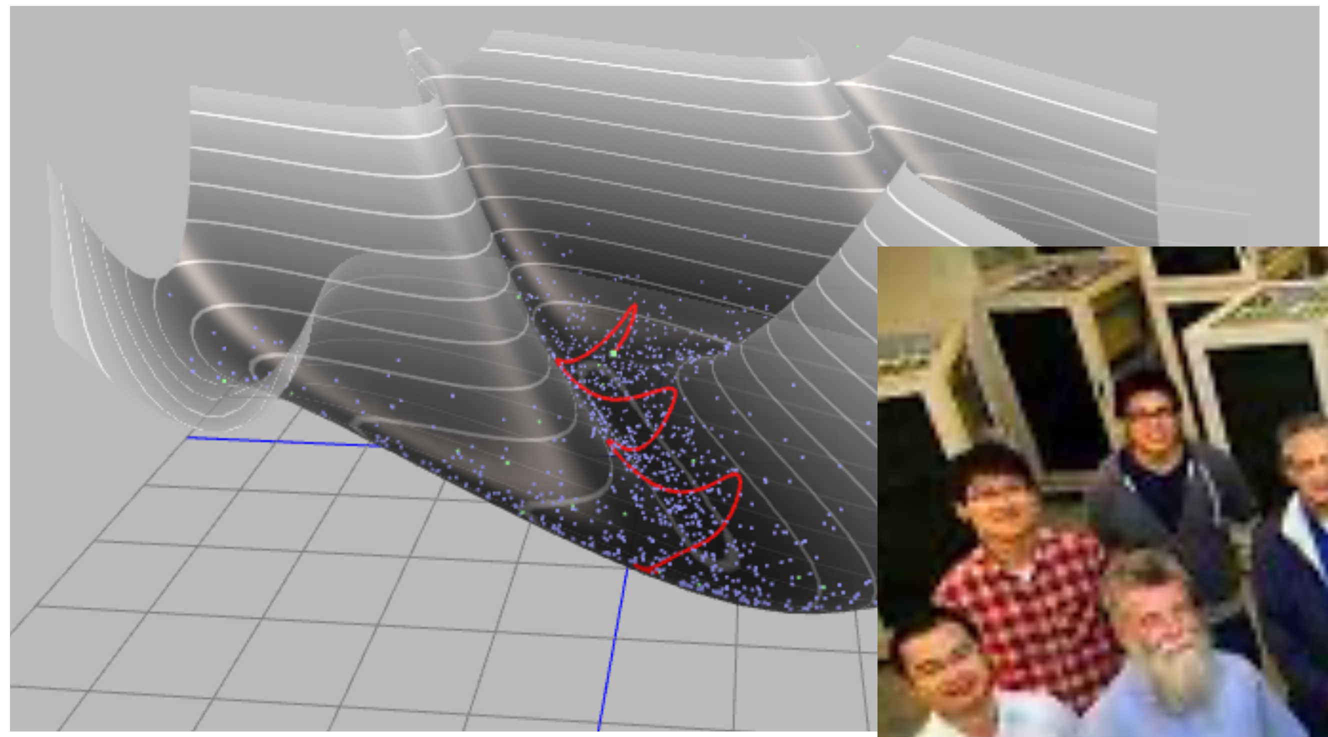
Long history of HEP driving innovation leading to interdisciplinary advances!

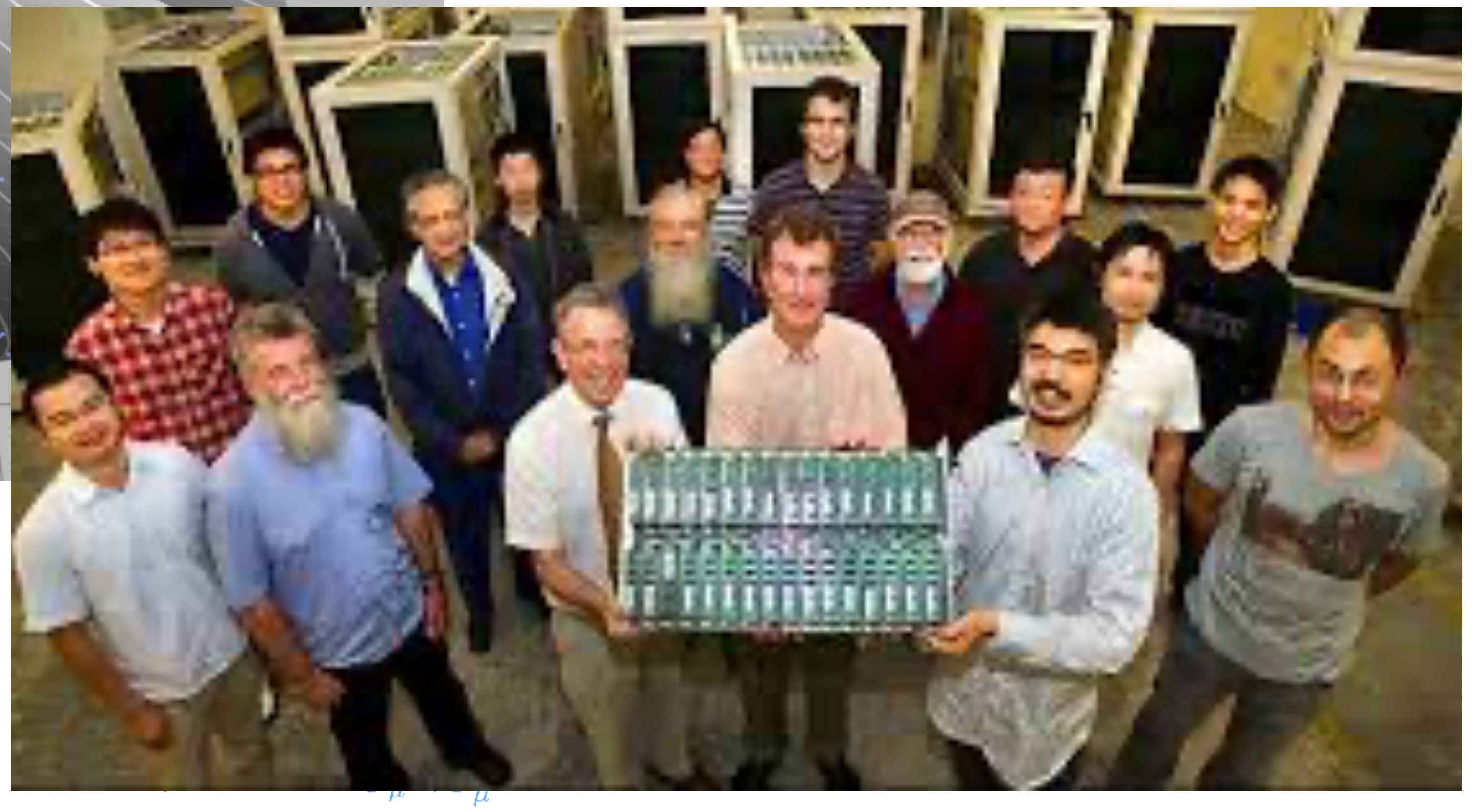
Theory case study: **Lattice QCD**

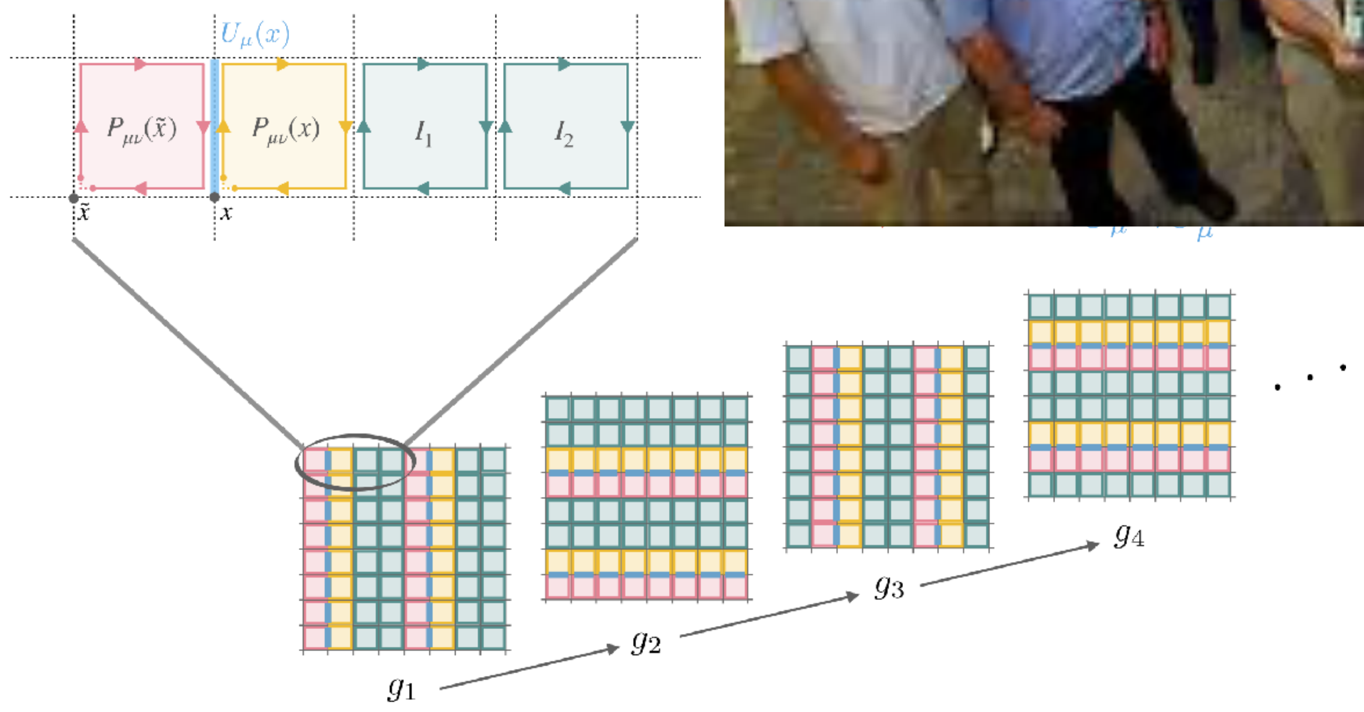
Numerical first-principles approach to non-perturbative QCD calculations

- Hamiltonian/Hybrid Monte Carlo (1980s)
- QCDOC ➡ Blue Gene supercomputers (2000s)
- Symmetry-equivariant ML sampling (2020s)

Same potential for technology transfer of future HEP-driven advancements in AI/ML!







HEP-inspired AI/ML can have broad impact

Long history of HEP driving innovation leading to interdisciplinary advances!

Theory case study: **Lattice QCD**

Numerical first-principles approach to non-perturbative QCD calculations

- Hamiltonian/Hybrid Monte Carlo (1980s)
- QCDOC ➡ Blue Gene supercomputers (2000s)
- Symmetry-equivariant ML sampling (2020s)

Same potential for technology transfer of future HEP-driven advancements in AI/ML!

- Markov-chain Monte Carlo sampling approach based on Hamiltonian dynamics
- Now a **widely-used workhorse algorithm for high-dimensional sampling problems**
- Applications across computational physics, chemistry, statistics (including ML)

HEP-inspired AI/ML can have broad impact

Long history of HEP driving innovation leading to interdisciplinary advances!

Theory case study: **Lattice QCD**

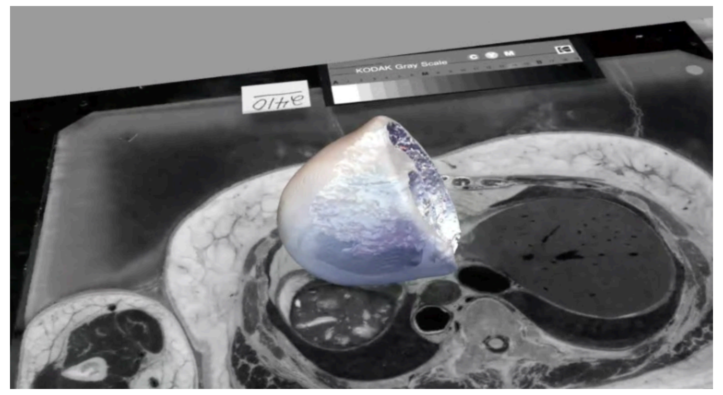
Numerical first-principles approach to non-perturbative QCD calculations

- Hamiltonian/Hybrid Monte Carlo (1980s)
- QCDOC ➡ Blue Gene supercomputers (2000s)
- Symmetry-equivariant ML sampling (2020s)

Same potential for technology transfer of future HEP-driven advancements in AI/ML!

- Universities/lab/industry (IBM) collaboration developed **massively parallel** architecture "QCD on a chip (QCDOC)" with **small footprint and power efficiency** that revolutionised HPC
- Pre-cursor of successful Blue Gene/L
- Enabled breakthrough applications in diverse areas e.g., tissue-level cardiac models

[IBM Cardioid Cardiac Modeling Project]



HEP-inspired AI/ML can have broad impact

Long history of HEP driving innovation leading to interdisciplinary advances!

Theory case study: **Lattice QCD**

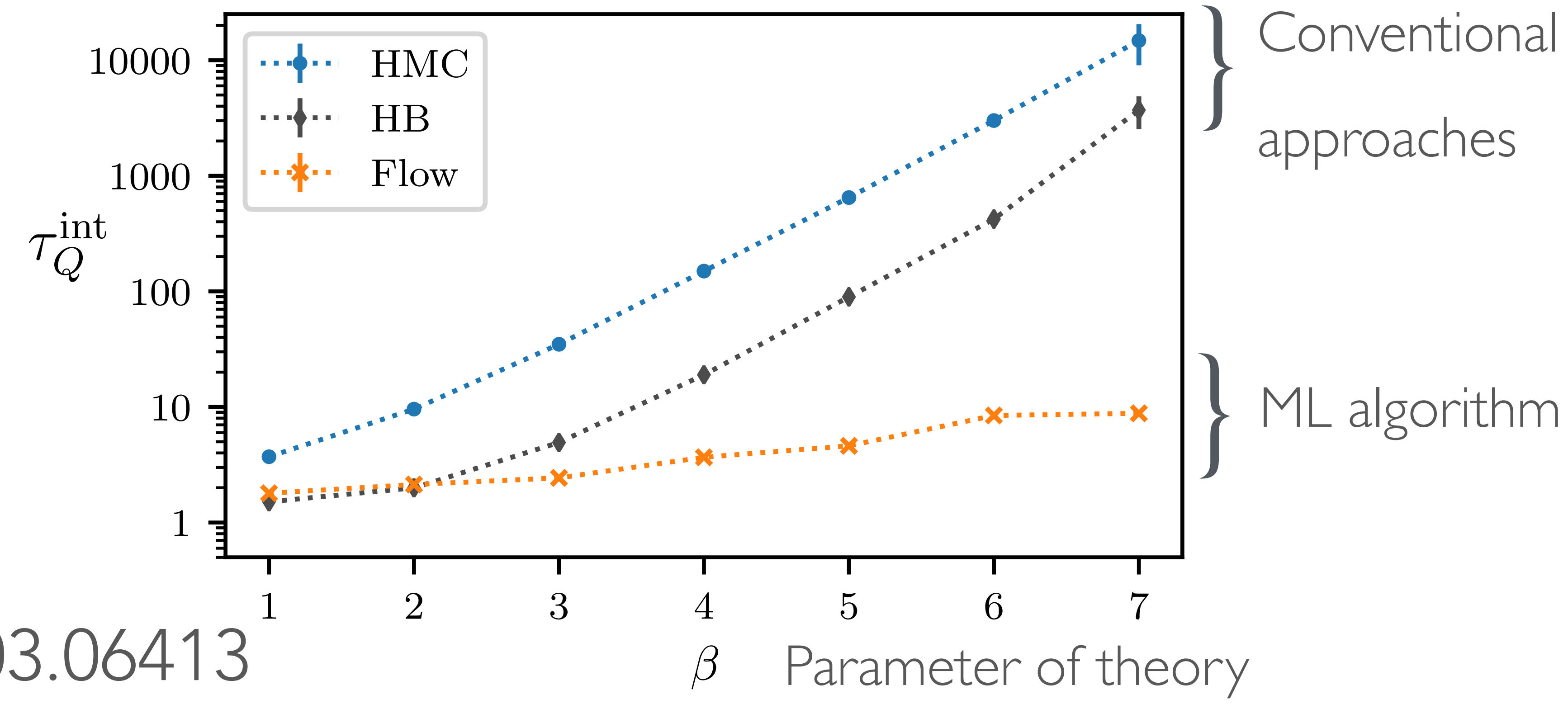
Numerical first-principles approach to non-perturbative QCD calculations

- Hamiltonian/Hybrid Monte Carlo (1980s)
- QCDOC ➡ Blue Gene supercomputers (2000s)
- Symmetry-equivariant ML sampling (2020s)

Same potential for technology transfer of future HEP-driven advancements in AI/ML!

- Generative models used to accelerate provably-exact sampling of lattice QCD gauge fields
- Exponential acceleration in proof-of-principle examples

Cost per independent sample



2003.06413

HEP-inspired AI/ML can have broad impact

Long history of HEP driving innovation leading to interdisciplinary advances!

Theory case study: **Lattice QCD**

Numerical first-principles approach to non-perturbative QCD calculations

- Hamiltonian/Hybrid Monte Carlo (1980s)
- QCDOC ➡ Blue Gene supercomputers (2000s)
- Symmetry-equivariant ML sampling (2020s)

Same potential for technology transfer of future HEP-driven advancements in AI/ML!

- Generative models used to accelerate provably-exact sampling of lattice QCD gauge fields
- Exponential acceleration in proof-of-principle examples
- Requires custom model architectures with physics built in
- Example of successful partnership with industry

What is needed to exploit AI/ML in HEP

Universities play, and will continue to play, a key role in AI/ML innovation

- **Workforce**
  - Advances are being driven by generation of young scientists trained at the physics/AI/ML intersection
  - Industry positions are appealing after graduation
    - Pipeline and long-term career prospects must be addressed
    - **Significant fraction of HEP AI/ML innovation is concentrated where there are junior researchers i.e., at universities**

Employment of New AI PhDs (% of Total) in North America by Sector, 2010–21
Source: CRA Taulbee Survey, 2022 | Chart: 2023 AI Index Report

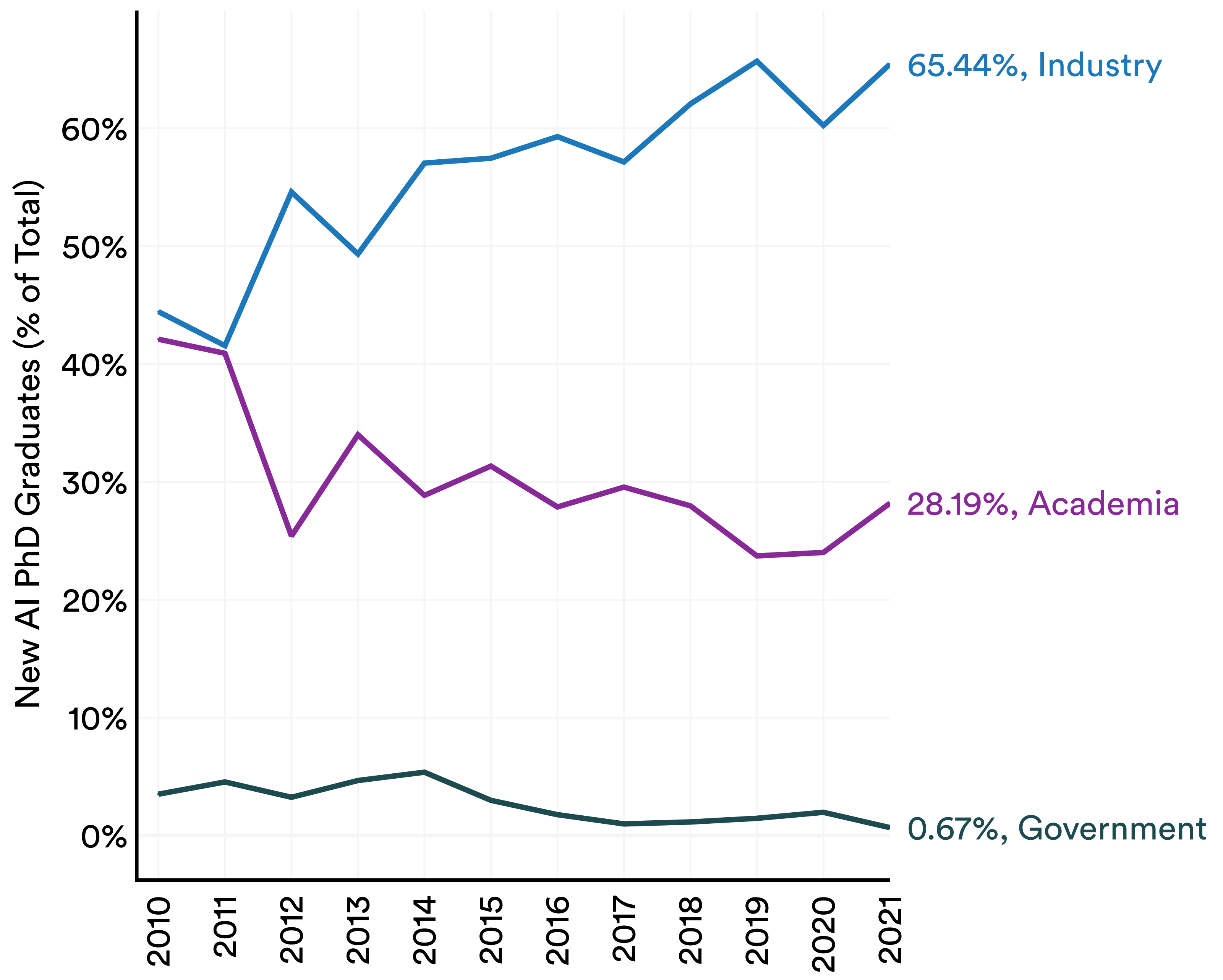


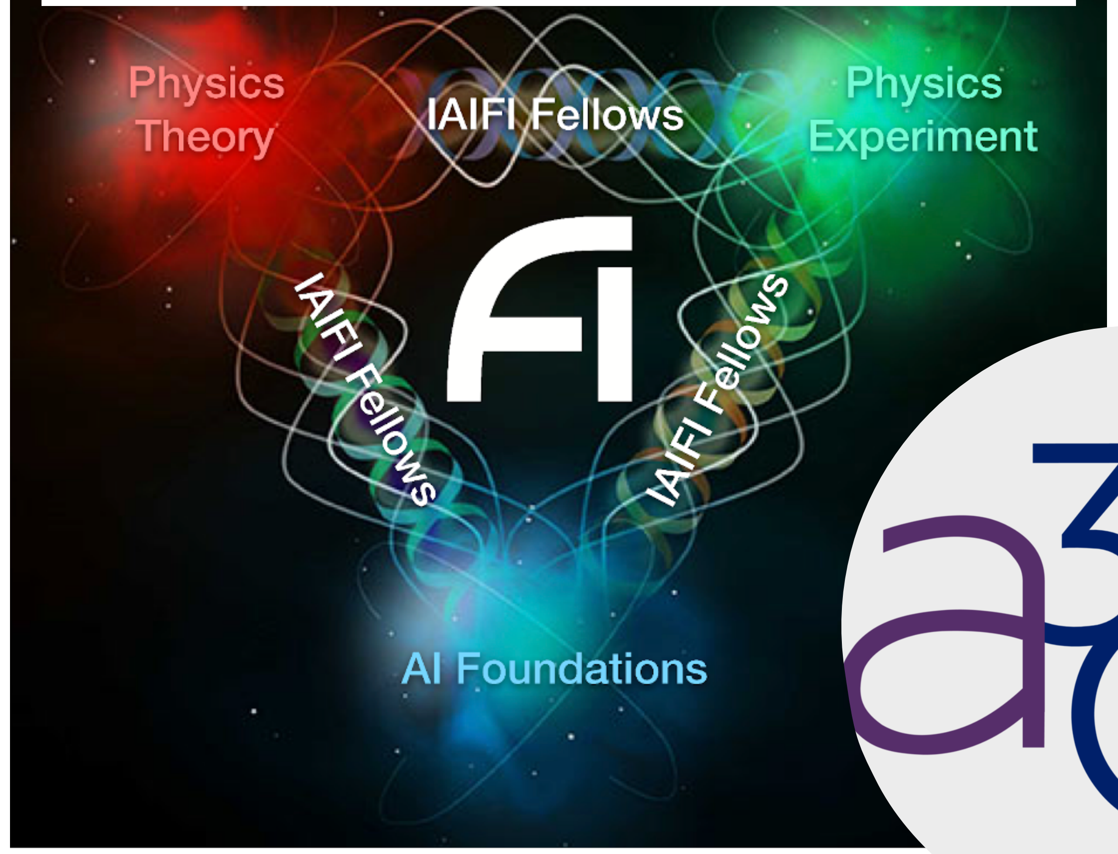
Figure 5.1.9

What is needed to exploit AI/ML in HEP

Deficiencies in current computing resources and allocation policies must be addressed

- **Computing resources**
  - Programs to develop and train talent at physics/AI/ML intersection are appearing, but **without significant computing resources for exploration and innovation**
  - Computing resources needed at all scales
  - Significant opportunity inequality when institutional university resources are a critical component for progress

The NSF Institute for Artificial Intelligence and Fundamental Interactions



Accelerated AI Algorithms for Data-Driven Discovery

Postbaccalaureate Research Fellows

The Eric and Wendy Schmidt AI in Science Postdoctoral Fellowship

Summary

Capitalising on **great potential** for transformative impact on HEP requires **targeted action**

- Transformational opportunities through both exploitation and “ground-up” ML/AI for HEP problems
    - **Demands support (people+hardware) for exploratory and developmental research at *both* universities and labs**
    - **Must train, retain, and capitalise on junior talent at physics/AI intersection**
      Collaborations with AI/ML “experts” external to physics community are necessary but not sufficient
- **Need support for AI/ML pipelines in HPC resource planning at all scales**

CompF Snowmass Report: 2210.05822
CompF3 Snowmass Report: 2209.07559