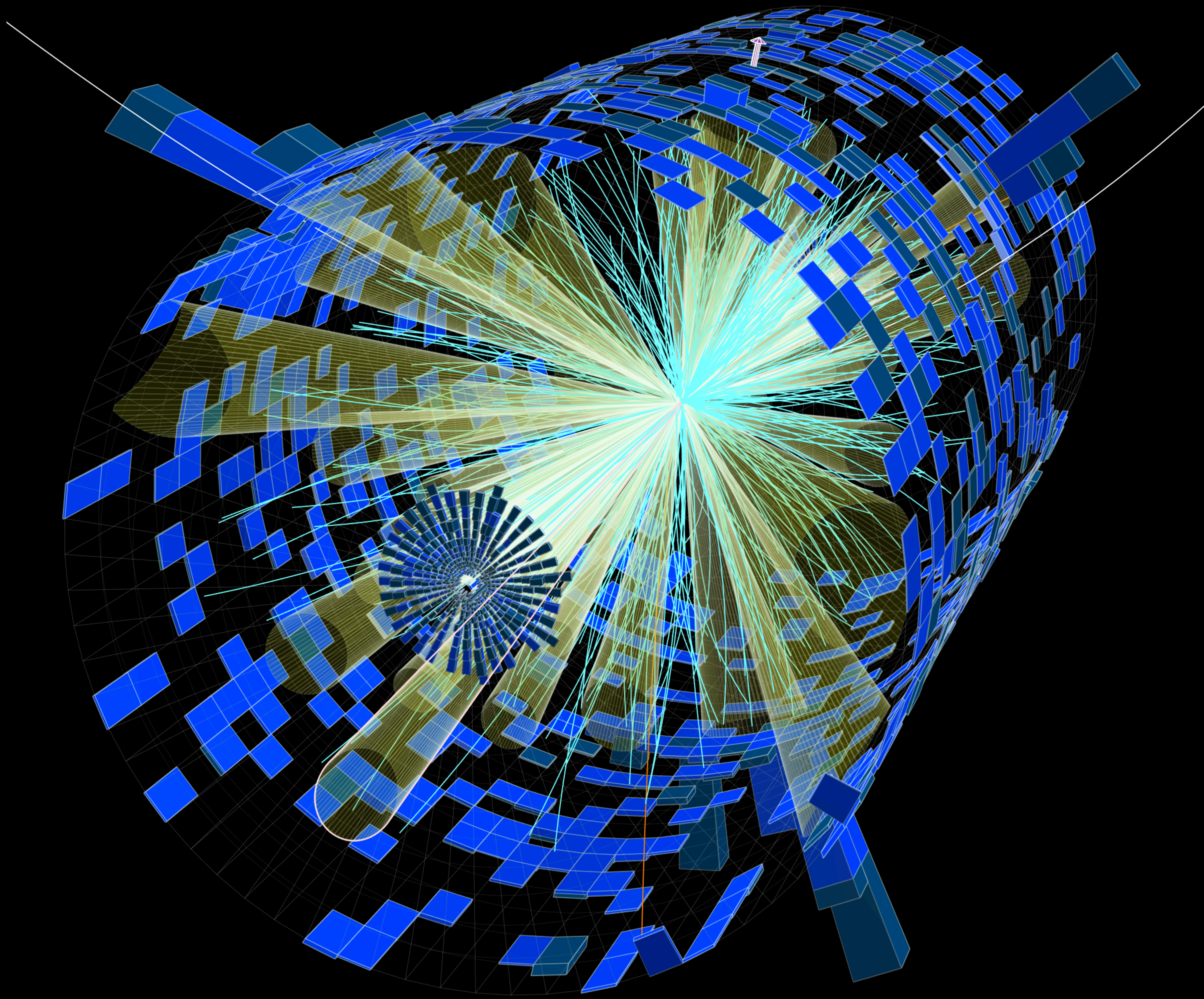
ARTIFICIAL INTELLIGENCE AND HEP



@KyleCranmer
University of Wisconsin-Madison
Data Science Institute, Physics, Statistics, Computer Science

Introduction

HEP has been using ML for decades (mainly for classification and regression), but what is happening recently is not just an improvement.

- We have qualitatively new capabilities
- What are they and why do they matter?

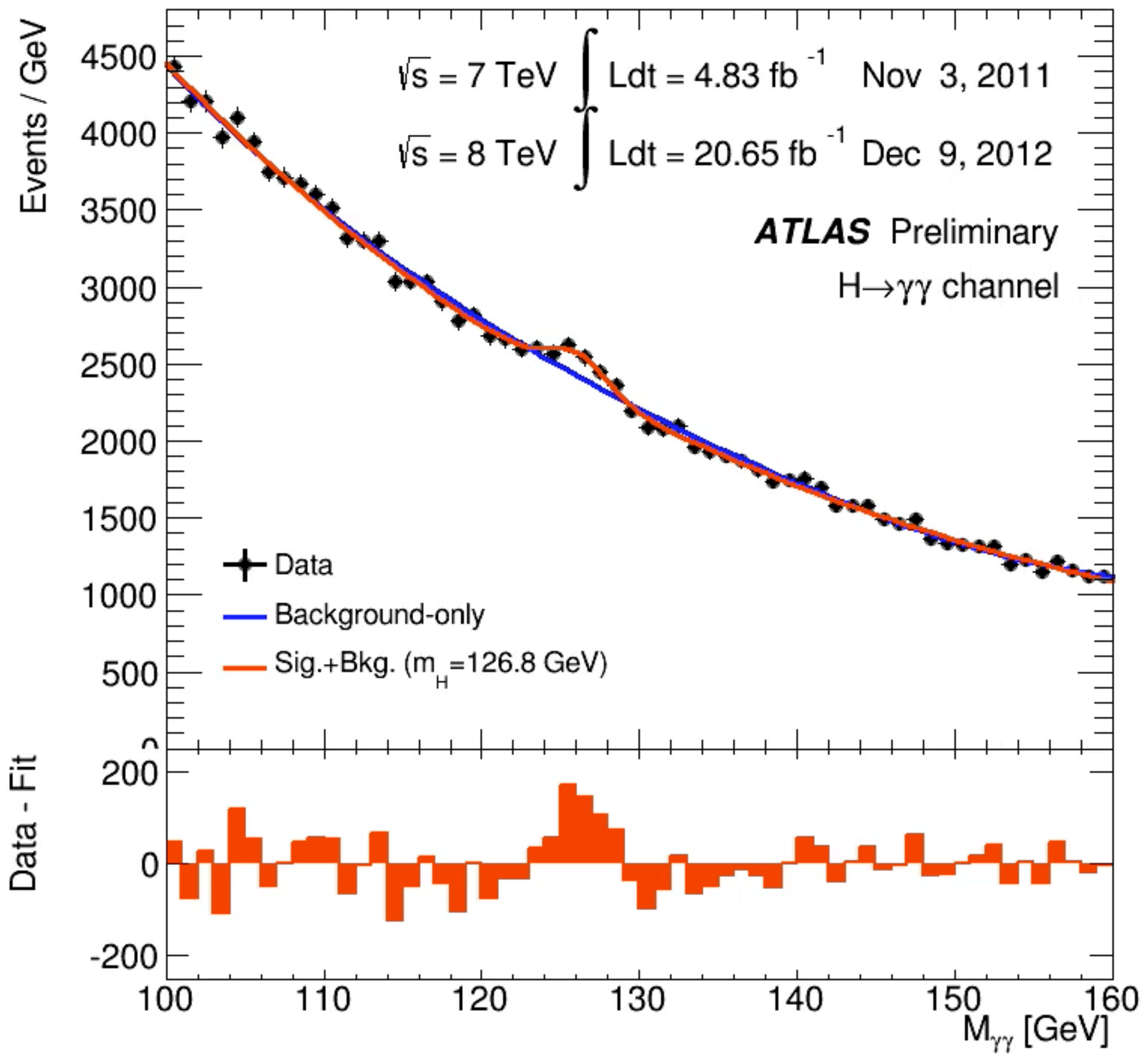
There is enormous hype around "AI" — is it physics?

- Focus on new capabilities, what they enable, & patterns of use

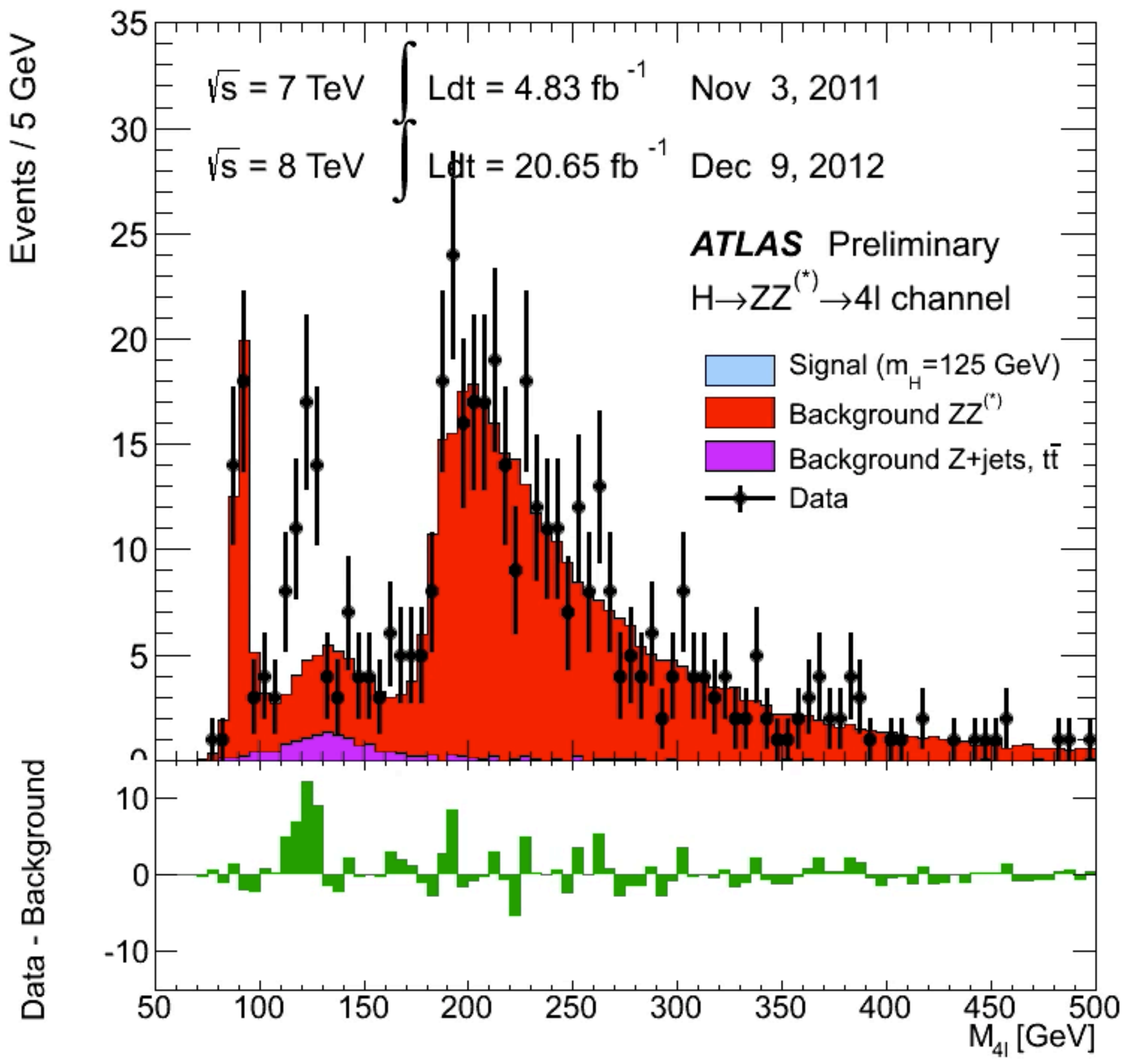
How does "AI4HEP" fit into broader trends around "AI4Science"?

- Important for arguing HEP's relevance to society, funding, etc.

July 2012

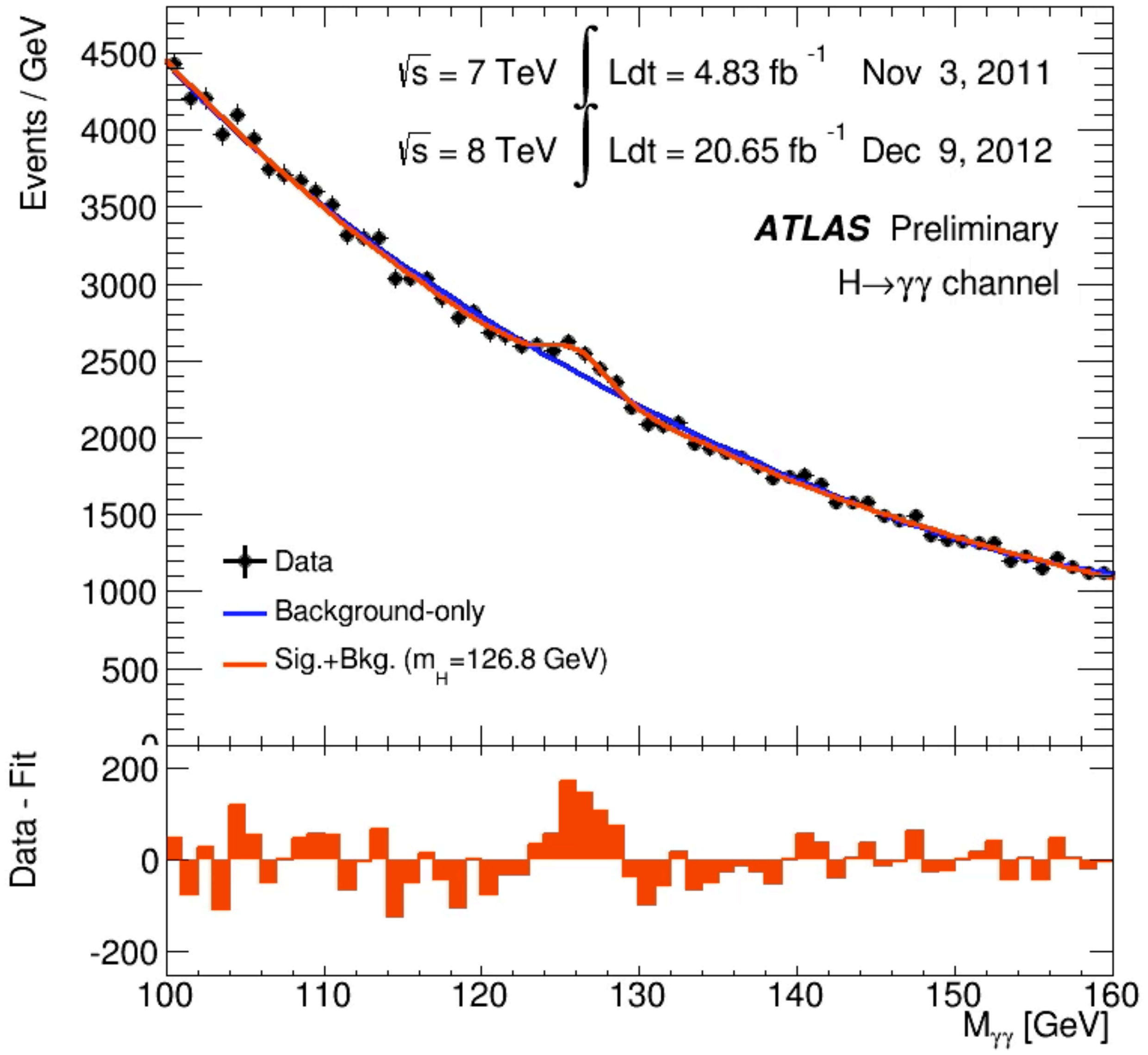


√s = 7 TeV ∫Ldt = 4.83 fb⁻¹ Nov 3, 2011
√s = 8 TeV ∫Ldt = 20.65 fb⁻¹ Dec 9, 2012
ATLAS Preliminary
H→γγ channel
Data
Background-only
Sig.+Bkg. (m_H=126.8 GeV)
Events / GeV
Data - Fit
M_γγ [GeV]

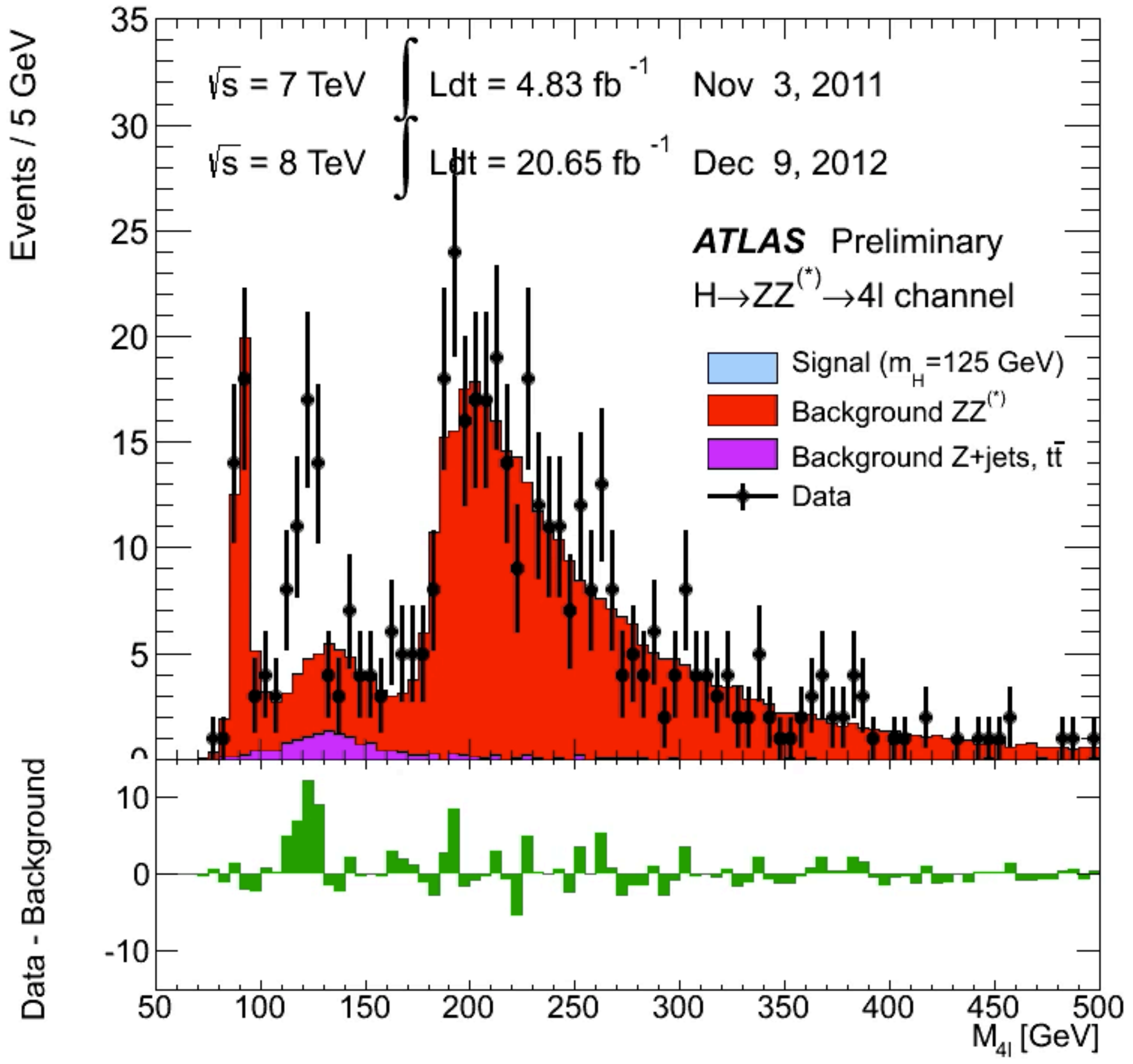


√s = 7 TeV ∫Ldt = 4.83 fb⁻¹ Nov 3, 2011
√s = 8 TeV ∫Ldt = 20.65 fb⁻¹ Dec 9, 2012
ATLAS Preliminary
H→ZZ(*)→4l channel
Signal (m_H=125 GeV)
Background ZZ(*)
Background Z+jets, tt̄
Data
Events / 5 GeV
Data - Background
M_4l [GeV]

July 2012

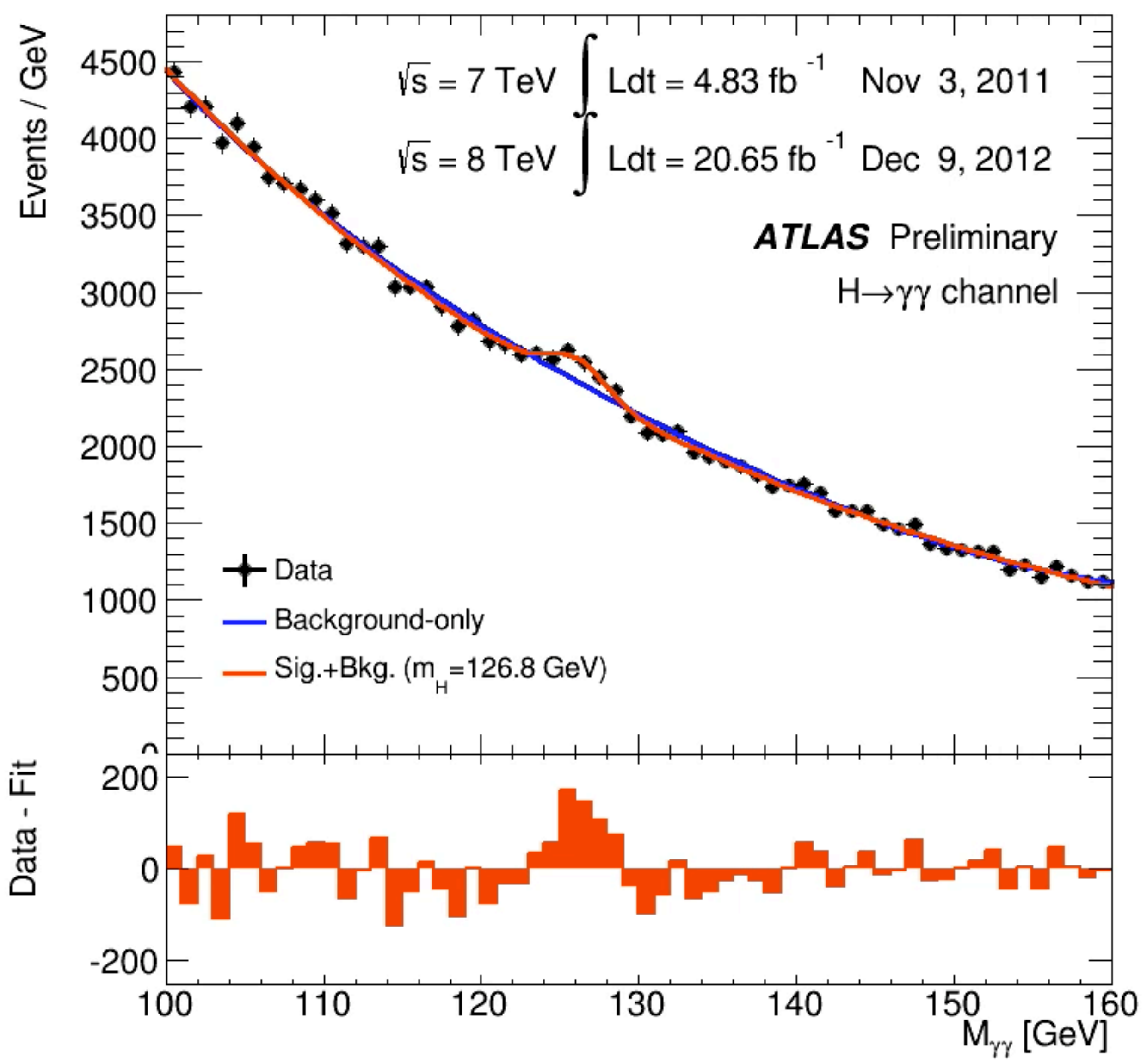


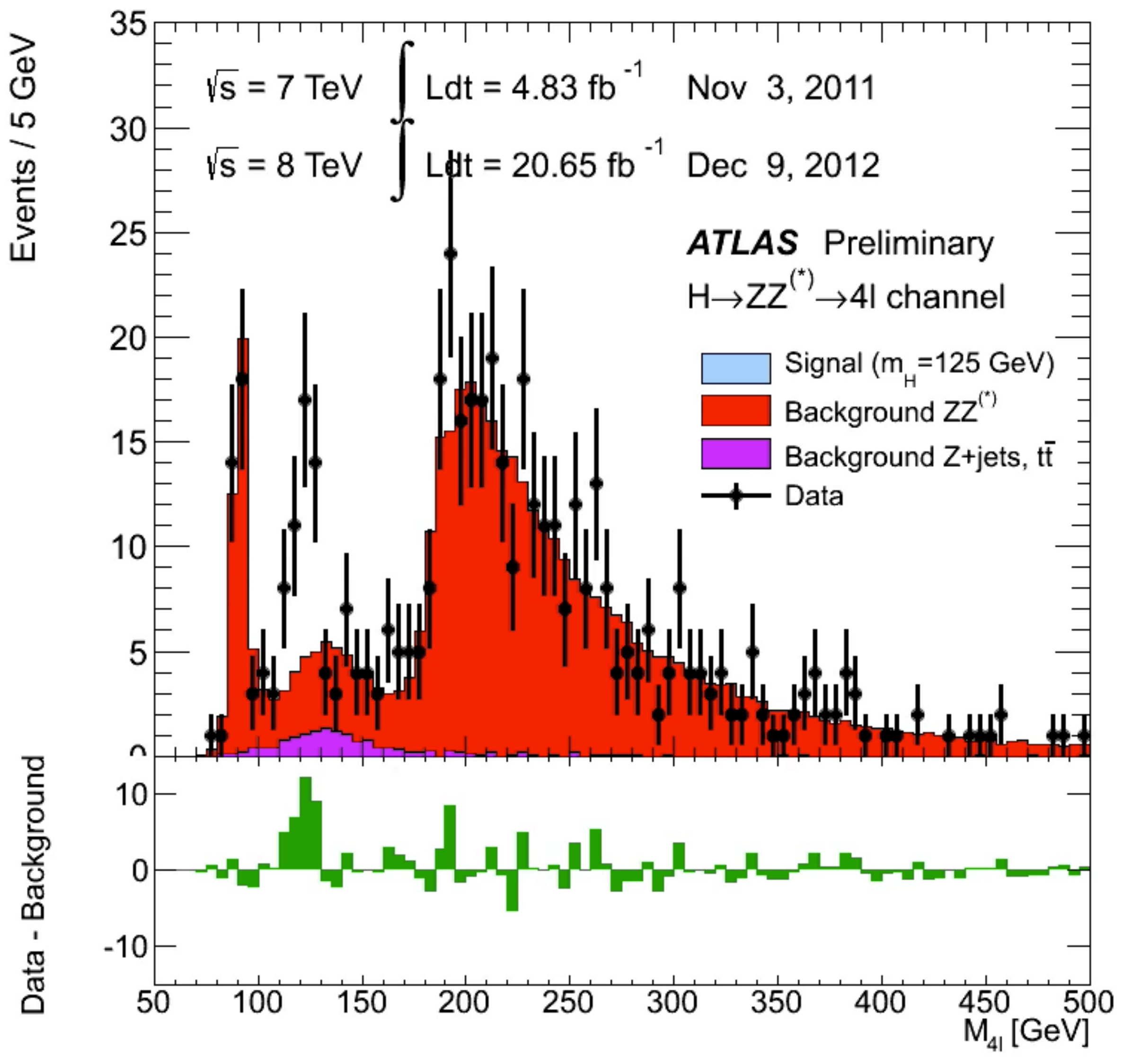
Events / GeV
√s = 7 TeV ∫Ldt = 4.83 fb⁻¹ Nov 3, 2011
√s = 8 TeV ∫Ldt = 20.65 fb⁻¹ Dec 9, 2012
ATLAS Preliminary
H→γγ channel
Data
Background-only
Sig.+Bkg. (m_H=126.8 GeV)
Data - Fit
M_γγ [GeV]

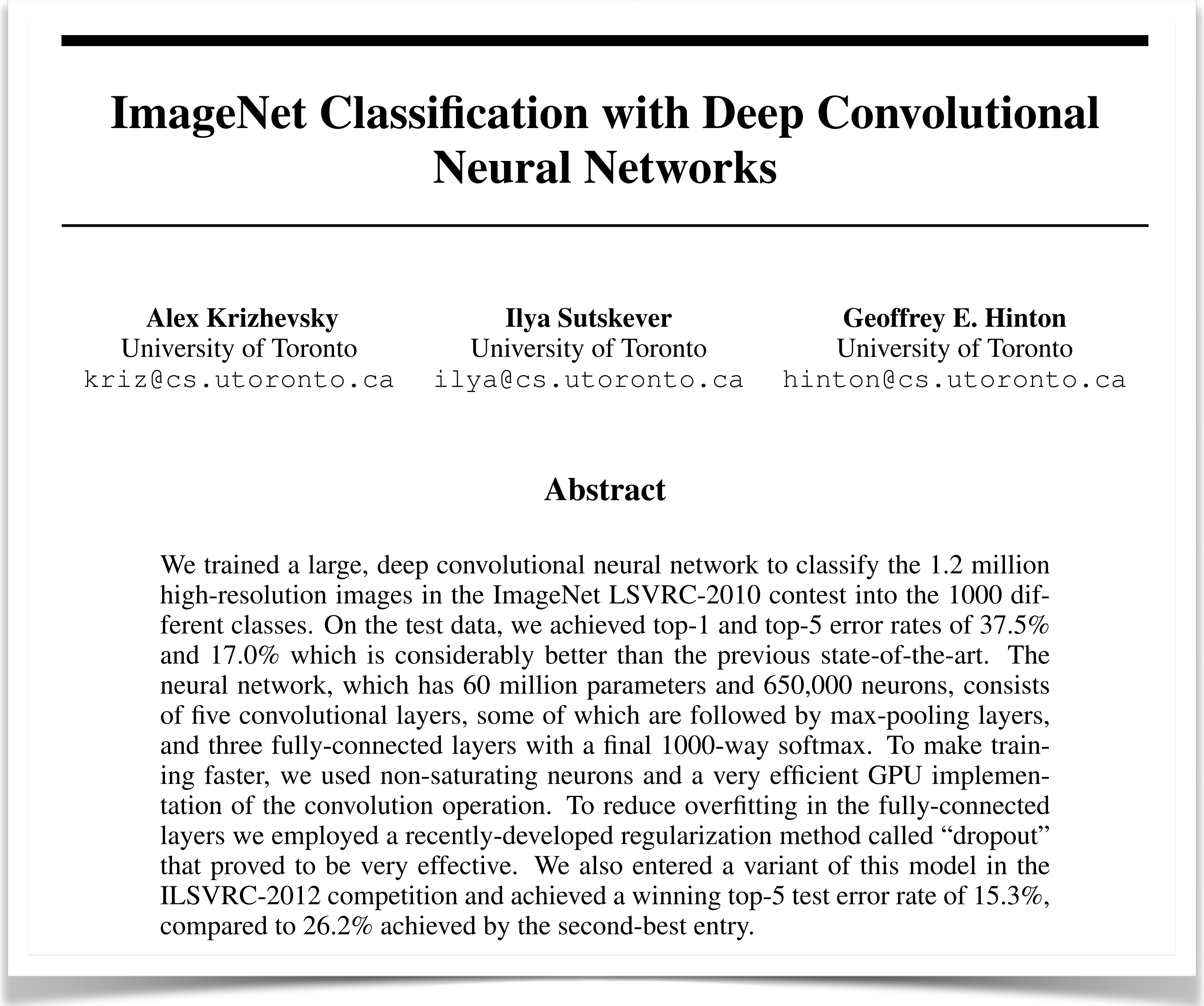


Events / 5 GeV
√s = 7 TeV ∫Ldt = 4.83 fb⁻¹ Nov 3, 2011
√s = 8 TeV ∫Ldt = 20.65 fb⁻¹ Dec 9, 2012
ATLAS Preliminary
H→ZZ(*)→4l channel
Signal (m_H=125 GeV)
Background ZZ(*)
Background Z+jets, tt̄
Data
Data - Background
M_4l [GeV]

July 2012







ImageNet Classification with Deep Convolutional Neural Networks

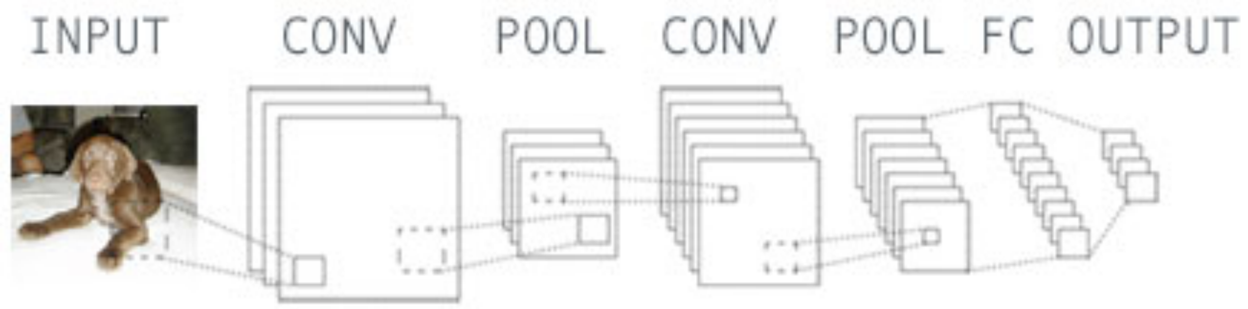
Alex Krizhevsky
University of Toronto
kriz@cs.utoronto.ca

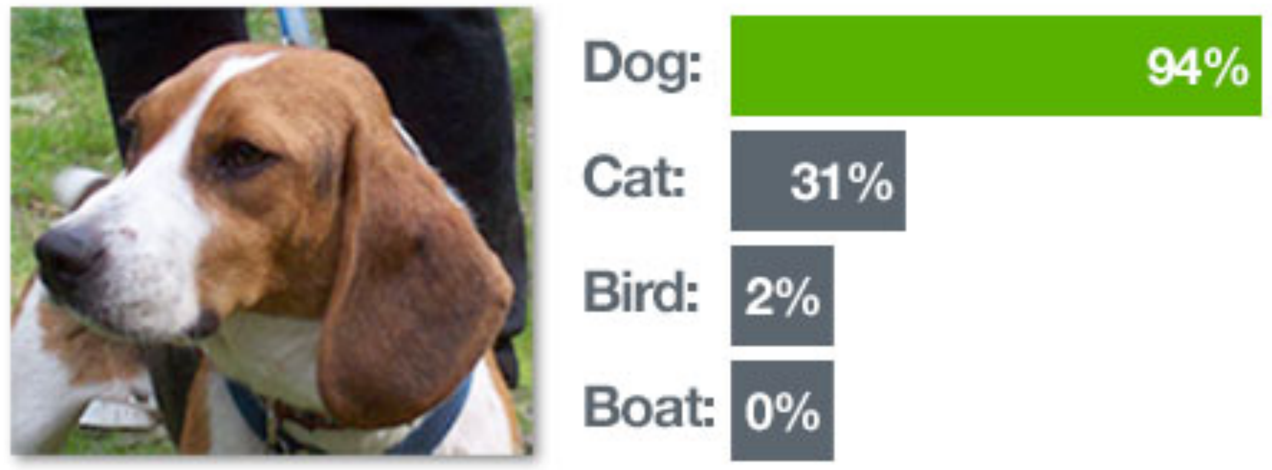
Ilya Sutskever
University of Toronto
ilya@cs.utoronto.ca

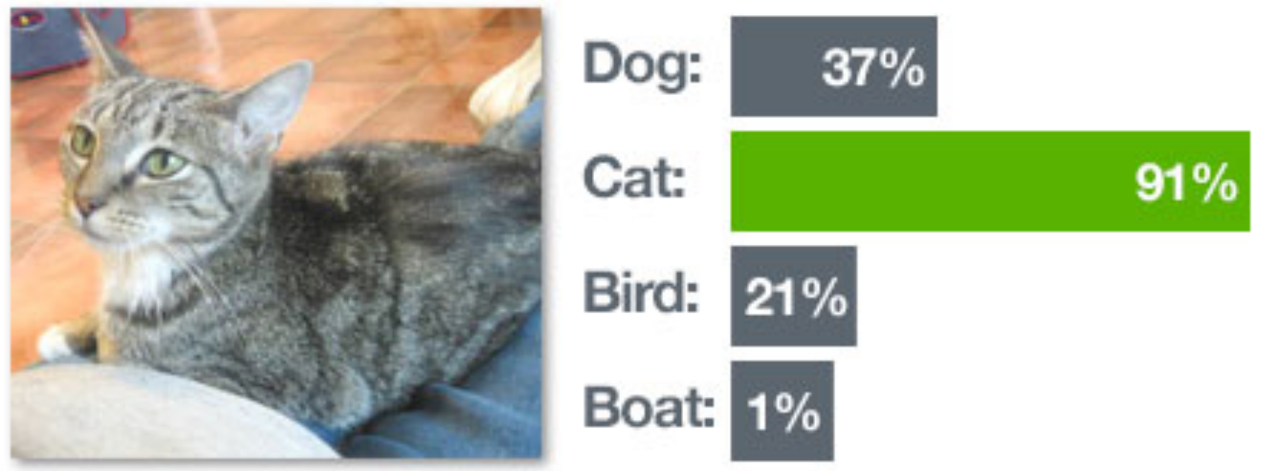
Geoffrey E. Hinton
University of Toronto
hinton@cs.utoronto.ca

Abstract

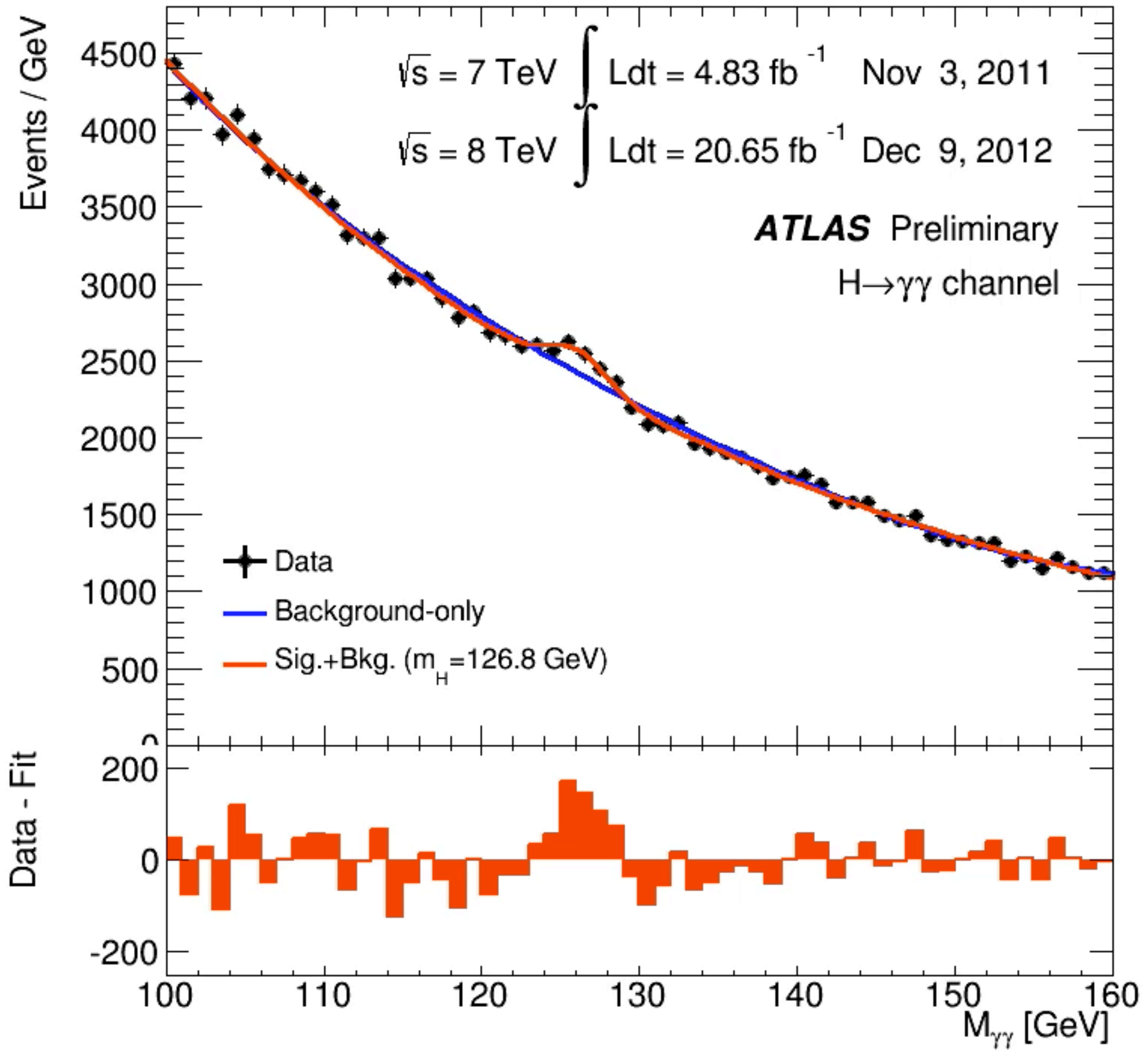
We trained a large, deep convolutional neural network to classify the 1.2 million high-resolution images in the ImageNet LSVRC-2010 contest into the 1000 different classes. On the test data, we achieved top-1 and top-5 error rates of 37.5% and 17.0% which is considerably better than the previous state-of-the-art. The neural network, which has 60 million parameters and 650,000 neurons, consists of five convolutional layers, some of which are followed by max-pooling layers, and three fully-connected layers with a final 1000-way softmax. To make training faster, we used non-saturating neurons and a very efficient GPU implementation of the convolution operation. To reduce overfitting in the fully-connected layers we employed a recently-developed regularization method called "dropout" that proved to be very effective. We also entered a variant of this model in the ILSVRC-2012 competition and achieved a winning top-5 test error rate of 15.3%, compared to 26.2% achieved by the second-best entry.

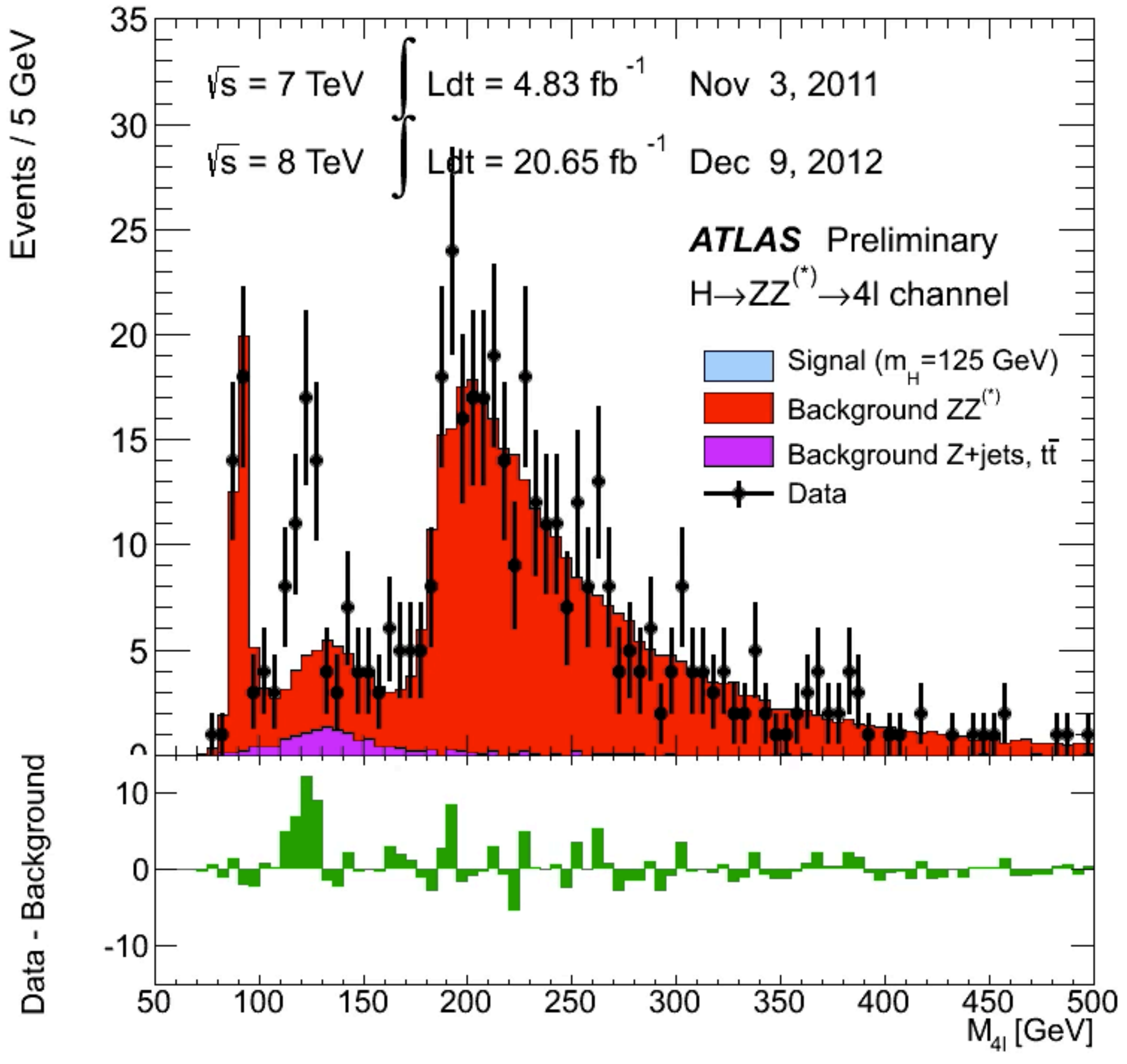


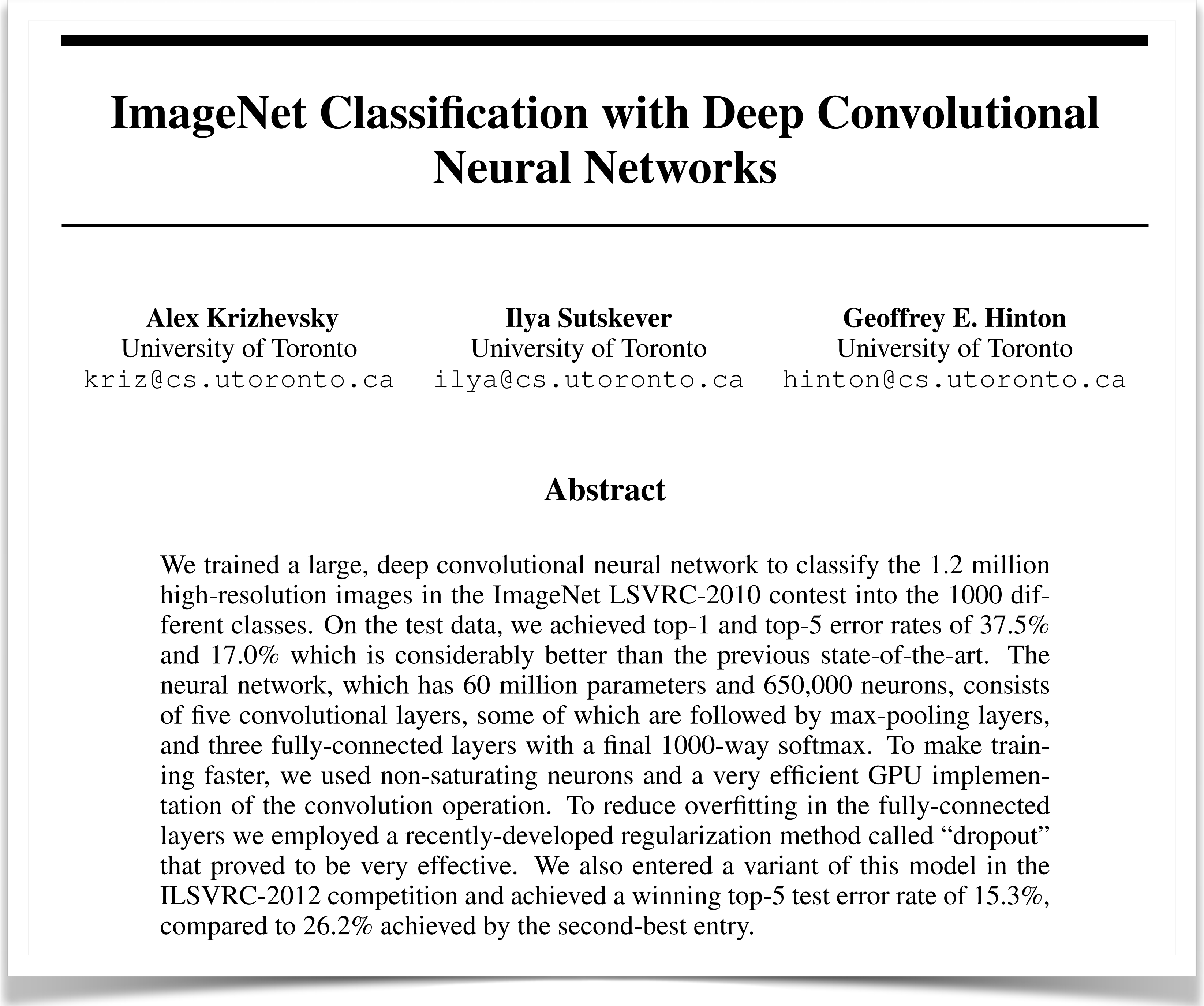




July 2012







ImageNet Classification with Deep Convolutional Neural Networks

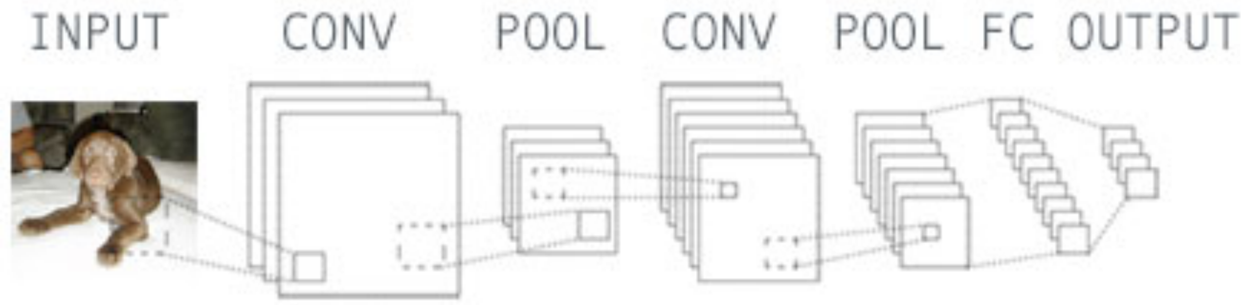
Alex Krizhevsky
University of Toronto
kriz@cs.utoronto.ca

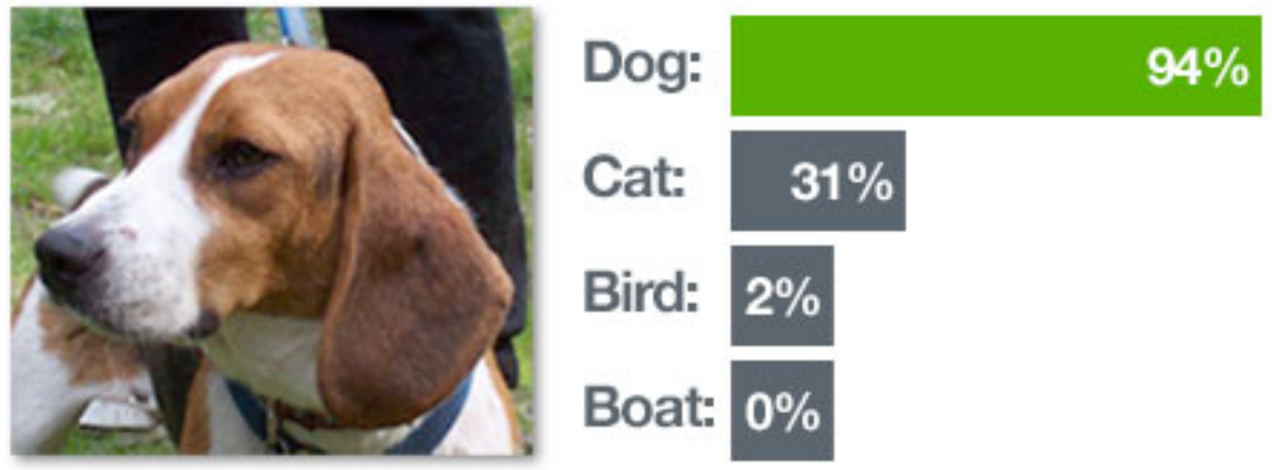
Ilya Sutskever
University of Toronto
ilya@cs.utoronto.ca

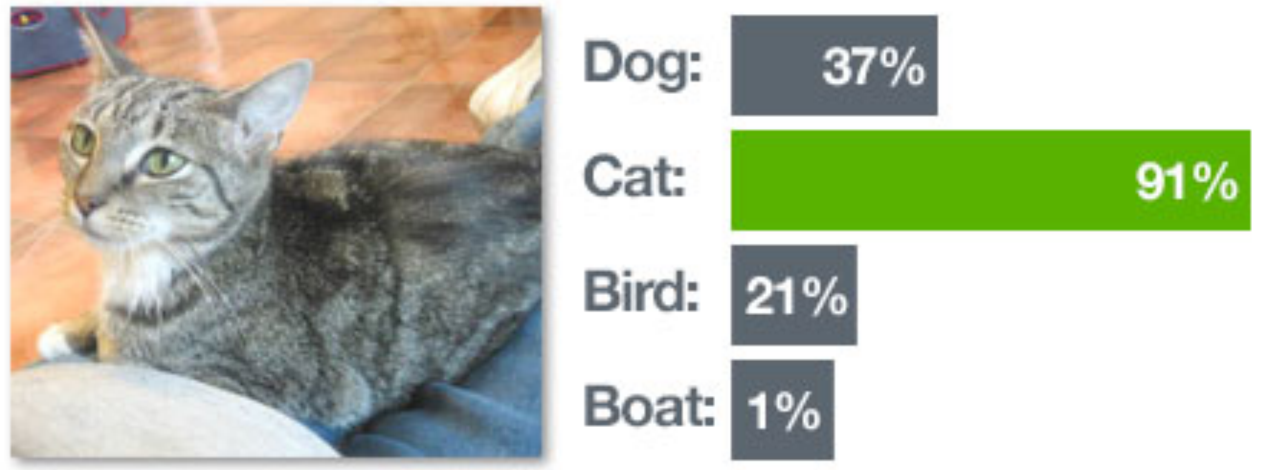
Geoffrey E. Hinton
University of Toronto
hinton@cs.utoronto.ca

Abstract

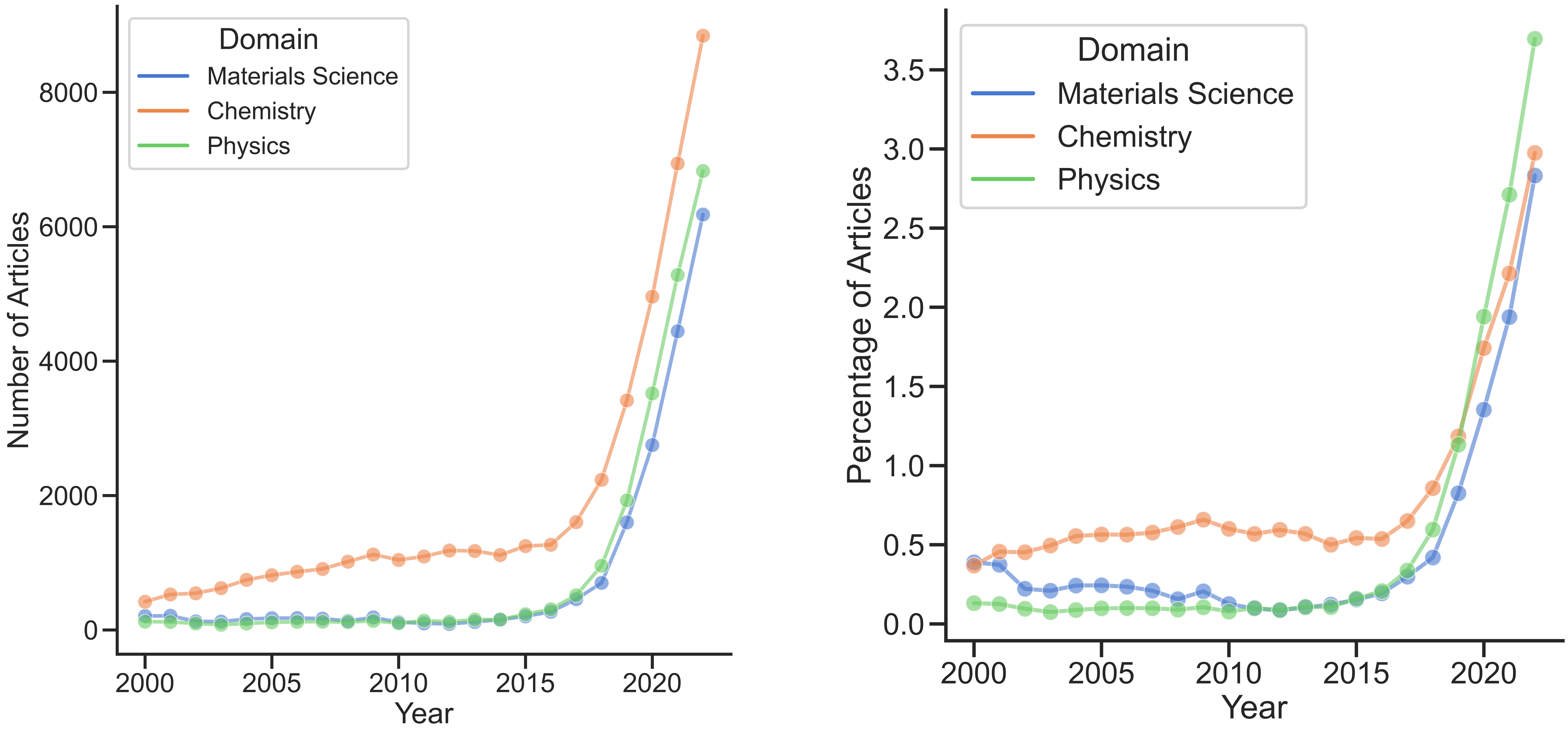
We trained a large, deep convolutional neural network to classify the 1.2 million high-resolution images in the ImageNet LSVRC-2010 contest into the 1000 different classes. On the test data, we achieved top-1 and top-5 error rates of 37.5% and 17.0% which is considerably better than the previous state-of-the-art. The neural network, which has 60 million parameters and 650,000 neurons, consists of five convolutional layers, some of which are followed by max-pooling layers, and three fully-connected layers with a final 1000-way softmax. To make training faster, we used non-saturating neurons and a very efficient GPU implementation of the convolution operation. To reduce overfitting in the fully-connected layers we employed a recently-developed regularization method called “dropout” that proved to be very effective. We also entered a variant of this model in the ILSVRC-2012 competition and achieved a winning top-5 test error rate of 15.3%, compared to 26.2% achieved by the second-best entry.







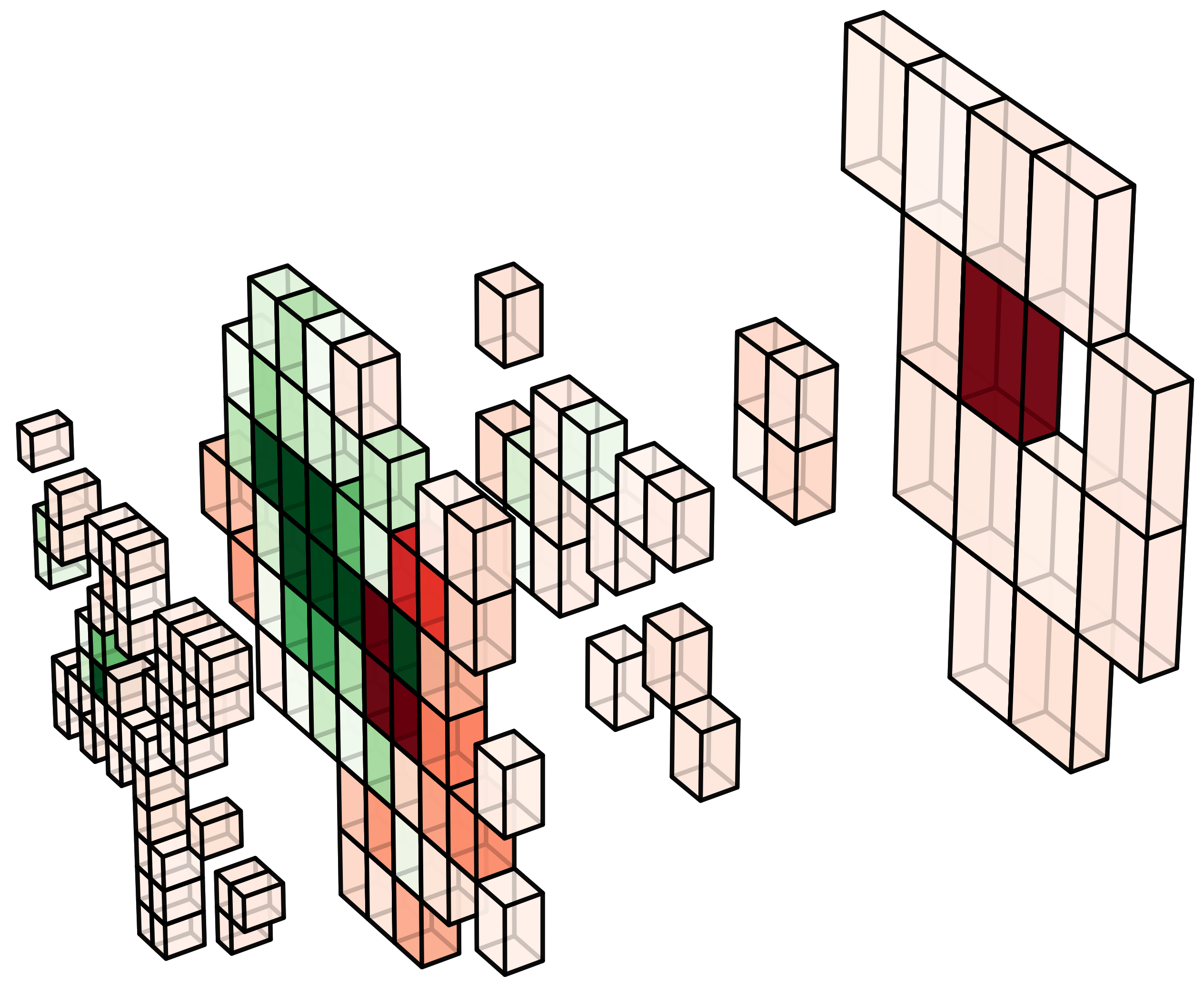
ML Publications in Science



Ben Blaiszik, “2021 AI/ML Publication Statistics and Charts”. Zenodo, Sep. 07, 2022. doi: 10.5281/zenodo.7057437.

Deep Learning's new capability

$$p_{T_1}, p_{T_2}, \eta_{T_1}, \eta_{T_2}, \Delta R_{jj}, \Delta\eta_{jj}, m_{jj}, \ldots$$



Tabular data

- "High-level features" / observables
  - angles, energies, inv. masses, …
- Fixed number of features
- ~ 5 - 30 observables

Richly-structured data

- Low-level objects
  - cells, clusters, tracks, …
- Variable number
- 100s-1000s of objects
  - Underlying geometry

A zoo of architectures

The ML4Jets Workshops

2017

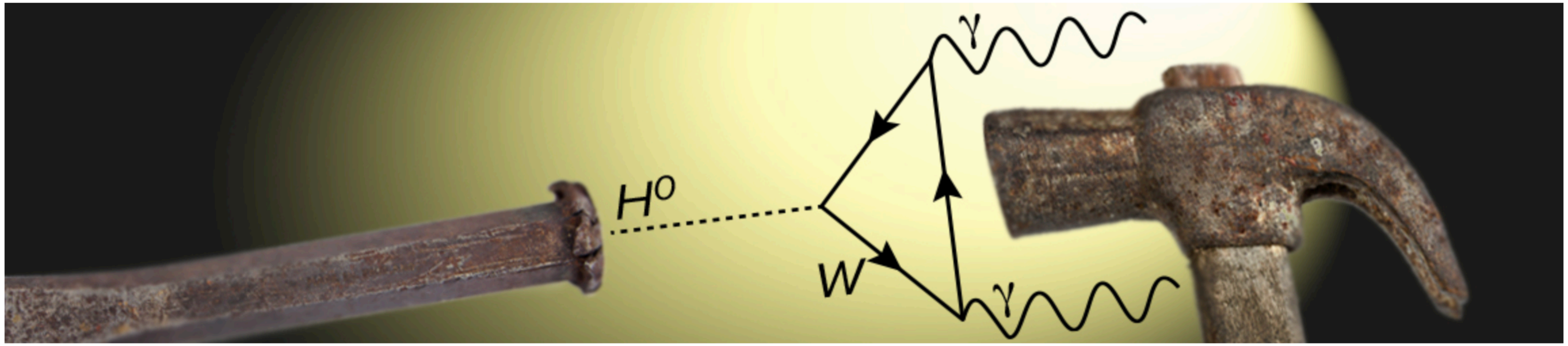
2020

2023

See agendas here: https://indico.cern.ch/event/1253794/

Hammers & Nails - Machine Learning & HEP

July 19-28, 2017 | Weizmann Institute of Science, Israel



Thanks Eilam and Toby!

2017

Topics include:

1. Generative models, high-dimensional density estimation, and likelihood-free inference
2. Sublinear-time pattern recognition and online learning
3. Domain adaptation and systematic uncertainty
4. Anomaly detection
5. Optimal experiment design and black box optimization
6. Generative Adversarial Network (GAN)
7. Geometric Deep Learning
8. U-Net

2019

Topics include:

1. Generative models, high-dimensional density estimation, and likelihood-free inference
2. Sublinear-time pattern recognition and online learning
3. Domain adaptation and systematic uncertainty
4. Anomaly detection
5. Optimal experiment design and black box optimization
6. Generative Adversarial Network (GAN)
7. Geometric Deep Learning

2022

Topics include:

1. Transformers, Attention, large language models (LLM), etc.
2. New types of generative models
3. Molecules (symmetries, graphs, generative models, etc.)
4. Uncertainty quantification and Bayesian NNs
5. Algorithmic reasoning
6. Optimal transport
7. Implicit layers
8. Variational inference / probabilistic reconstruction
9. Self-supervised learning

https://conferences.weizmann.ac.il/SRitp/Aug2022/hammers-nails-2022-0

Machine Learning and the Physical Sciences

Workshop at the 39th conference on Neural Information Processing Systems (NeurIPS)

December 6 or 7, 2025

A new journal focusing on ML for Physics

MACHINE LEARNING
Science and Technology

PAPER • OPEN ACCESS
Stochastic black-box optimization using multi-fidelity score function estimator
Atul Agrawal*, Kislaya Ravi, Phaedon-Stelios Koutsourelakis and Hans-Joachim Bungartz
Published 31 January 2025 • © 2025 The Author(s). Published by IOP Publishing Ltd
Machine Learning: Science and Technology, Volume 6, Number 1
Focus on ML and the Physical Sciences
Citation Atul Agrawal *et al* 2025 *Mach. Learn.: Sci. Technol.* **6** 015024
DOI 10.1088/2632-2153/ad8e2b

OPEN ACCESS
Comparing AI versus optimization workflows for simulation-based inference of spatial-stochastic systems
Michael Alexander Ramirez Sierra* and Thomas R Sokolowski
Published 14 February 2025 • © 2025 The Author(s). Published by IOP Publishing Ltd
Machine Learning: Science and Technology, Volume 6, Number 1
Citation Michael Alexander Ramirez Sierra and Thomas R Sokolowski 2025 *Mach. Learn.: Sci. Technol.* **6** 010502
DOI 10.1088/2632-2153/ada0a3

PAPER • OPEN ACCESS
Simulation-based inference with approximately correct parameters via maximum entropy
Rainier Barrett, Mehrad Ansari, Gourab Ghoshal and Andrew D White
Published 27 April 2022 • © 2022 The Author(s). Published by IOP Publishing Ltd
Machine Learning: Science and Technology, Volume 3, Number 2
Citation Rainier Barrett *et al* 2022 *Mach. Learn.: Sci. Technol.* **3** 025006
DOI 10.1088/2632-2153/ac6286

LETTER • OPEN ACCESS
DIGS: deep inference of galaxy spectra with neural posterior estimation
Gourav Khullar, Brian Nord, Aleksandra Ćiprijanović, Jason Poh and Fei Xu
Published 28 December 2022 • © 2022 The Author(s). Published by IOP Publishing Ltd
Machine Learning: Science and Technology, Volume 3, Number 4
Citation Gourav Khullar *et al* 2022 *Mach. Learn.: Sci. Technol.* **3** 04LT04
DOI 10.1088/2632-2153/ac98f4

PAPER • OPEN ACCESS
Evidence Networks: simple losses for fast, amortized, neural Bayesian model comparison
Niall Jeffrey and Benjamin D Wandelt
Published 17 January 2024 • © 2024 The Author(s). Published by IOP Publishing Ltd
Machine Learning: Science and Technology, Volume 5, Number 1
Citation Niall Jeffrey and Benjamin D Wandelt 2024 *Mach. Learn.: Sci. Technol.* **5** 015008
DOI 10.1088/2632-2153/ad1a4d

PAPER • OPEN ACCESS
Simulation-based inference on virtual brain models of disorders
Meysam Hashemi, Abolfazl Ziaeemehr, Marmaduke M Woodman, Jan Fousek, Spase Petkoski and Viktor K Jirsa
Published 19 July 2024 • © 2024 The Author(s). Published by IOP Publishing Ltd
Machine Learning: Science and Technology, Volume 5, Number 3
Focus on Explainable Machine Learning in Sciences
Citation Meysam Hashemi *et al* 2024 *Mach. Learn.: Sci. Technol.* **5** 035019
DOI 10.1088/2632-2153/ad6230

PAPER • OPEN ACCESS
Importance nested sampling with normalising flows
Michael J Williams, John Veitch and Chris Messenger
Published 25 July 2023 • © 2023 The Author(s). Published by IOP Publishing Ltd
Machine Learning: Science and Technology, Volume 4, Number 3
Citation Michael J Williams *et al* 2023 *Mach. Learn.: Sci. Technol.* **4** 035011
DOI 10.1088/2632-2153/acd5aa

PAPER • OPEN ACCESS
Multi-fidelity Gaussian process surrogate modeling for regression problems in physics
Kislaya Ravi*, Vladyslav Fediukov*, Felix Dietrich, Tobias Neckel, Fabian Buse, Michael Bergmann and Hans-Joachim Bungartz
Published 15 October 2024 • © 2024 The Author(s). Published by IOP Publishing Ltd
Machine Learning: Science and Technology, Volume 5, Number 4
Focus on ML and the Physical Sciences
Citation Kislaya Ravi *et al* 2024 *Mach. Learn.: Sci. Technol.* **5** 045015
DOI 10.1088/2632-2153/ad7ad5

PAPER • OPEN ACCESS
Efficient Bayesian inference using physics-informed invertible neural networks for inverse problems
Xiaofei Guan, Xintong Wang, Hao Wu, Zihao Yang and Peng Yu
Published 23 July 2024 • © 2024 The Author(s). Published by IOP Publishing Ltd
Machine Learning: Science and Technology, Volume 5, Number 3
Focus on Generative AI in Science
Citation Xiaofei Guan *et al* 2024 *Mach. Learn.: Sci. Technol.* **5** 035026
DOI 10.1088/2632-2153/ad5f74

PAPER • OPEN ACCESS
`DiffLense`: a conditional diffusion model for super-resolution of gravitational lensing data
Pranath Reddy*, Michael W Toomey, Hanna Parul and Sergei Gleyzer
Published 19 September 2024 • © 2024 The Author(s). Published by IOP Publishing Ltd
Machine Learning: Science and Technology, Volume 5, Number 3
Focus on ML and the Physical Sciences
Citation Pranath Reddy *et al* 2024 *Mach. Learn.: Sci. Technol.* **5** 035076
DOI 10.1088/2632-2153/ad76f8

Disclaimer: I'm Editor-in-Chief http://iopscience.iop.org/mlst

Insight of data generating process informs inductive bias on architecture

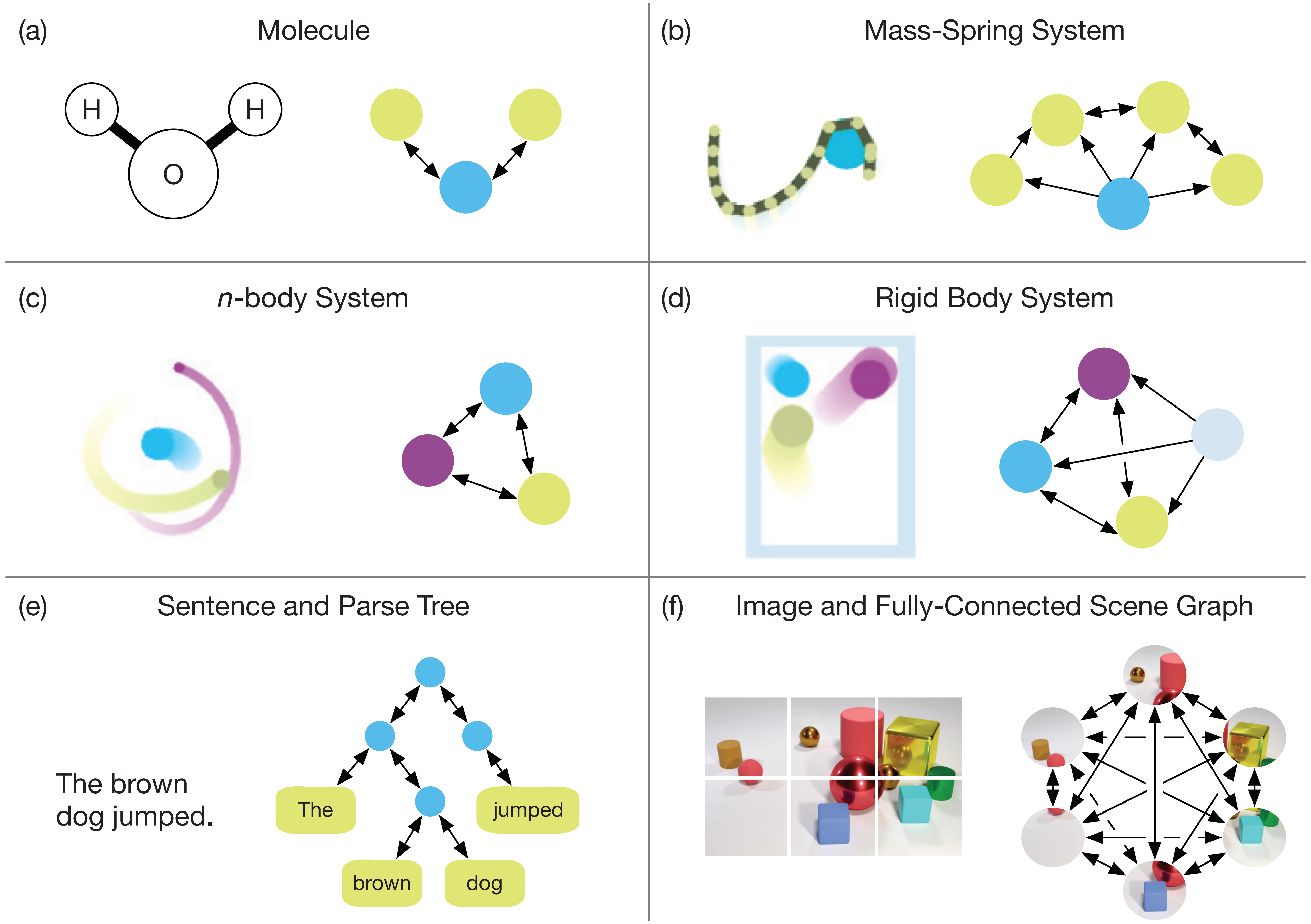


Image credit: Battaglia, et. al. arXiv:1806.01261

Inductive Bias

Compositionality

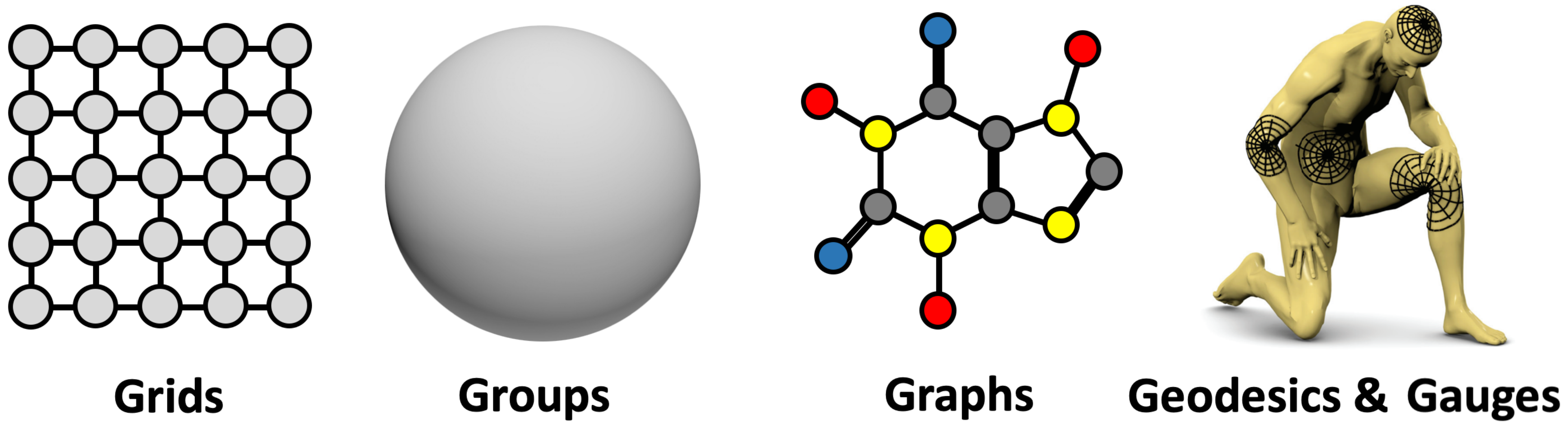
Relationships

Symmetry

Causality

Separation

HEP's problems, experience, and contributions were recognized by the AI/ML community

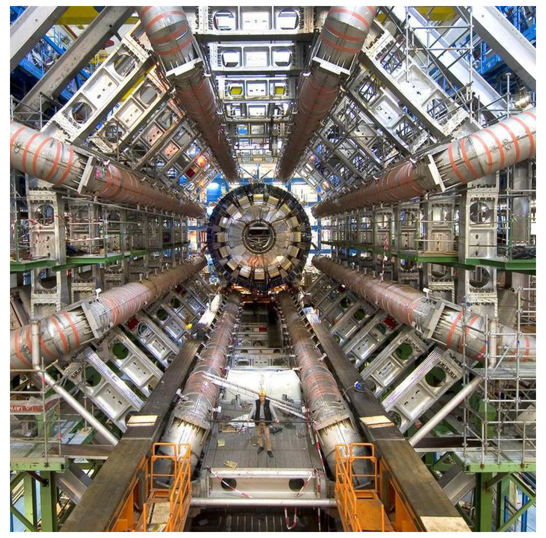


Geometric Deep Learning
Grids, Groups, Graphs,
Geodesics, and Gauges

Michael M. Bronstein[1], Joan Bruna[2], Taco Cohen[3], Petar Veličković[4]

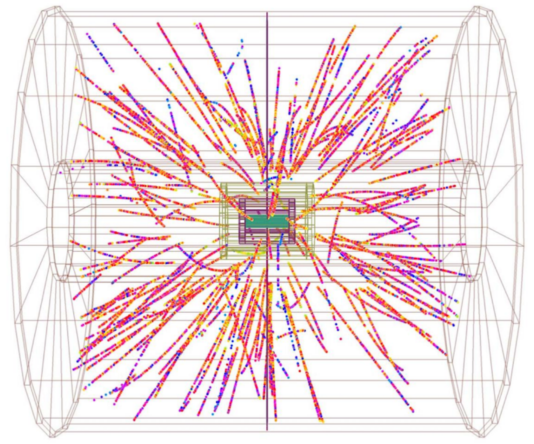
May 4, 2021

arXiv:2104.13478


Part of the Large Hadron Collider detectors.

Particle physics and astrophysics High energy physicists were perhaps among the first domain experts in the field of natural sciences to embrace the new shiny tool, graph neural networks. In a recent review paper, Shlomi et al. (2020) note that machine learning has historically been heavily used in particle physics experiments, either to learn complicated inverse functions allowing to infer the underlying physics process from the information measured in the detector, or to perform classification and regression tasks. For the latter, it was often necessary to force the data into an unnatural representation such as grid, in order to be able to used standard deep learning architectures such as CNN. Yet, many problems in physics involve data in the form of unordered sets with rich relations and interactions, which can be naturally represented as graphs.

One important application in high-energy physics is the reconstruction and classification of *particle jets* – sprays of stable particles arising from multiple successive interaction and decays of particles originating from a single initial event. In the Large Hardon Collider, the largest and best-known particle accelerator built at CERN, such jet are the result of collisions of protons at nearly the speed of light. These collisions produce massive particles, such as the long though-for Higgs boson or the top quark. The identification and classification of collision events is of crucial importance, as it might provide experimental evidence to the existence of new particles.


Example of a particle jet.

Multiple Geometric Deep Learning approaches have recently been proposed for particle jet classification task, e.g. by Komiske et al. (2019) and Qu and Gouskos (2019), based on DeepSet and Dynamic Graph CNN architectures, respectively. More recently, there has also been interest in developing specialsed architectures derived from physics consideration and incorporating inductive biases consistent with Hamiltonian or Lagrangian mechanics (see e.g. Sanchez-Gonzalez et al. (2019); Cranmer et al. (2020)), equivariant to the Lorentz group (a fundamental symmetry of space and time in physics) (Bogatskiy et al., 2020), or even incorporating symbolic reasoning (Cranmer et al., 2019) and capable of learning physical laws from data. Such approaches are more interpretable (and thus considered more 'trustworthy' by domain experts) and also offer better generalisation.

Besides particle accelerators, particle detectors are now being used by astrophysicist for *multi-messenger astronomy* – a new way of coordinated observation of disparate signals, such as electromagnetic radiation, gravitational waves, and neutrinos, coming from the same source. Neutrino astronomy is

Rapid progress

arXiv:1902.09914

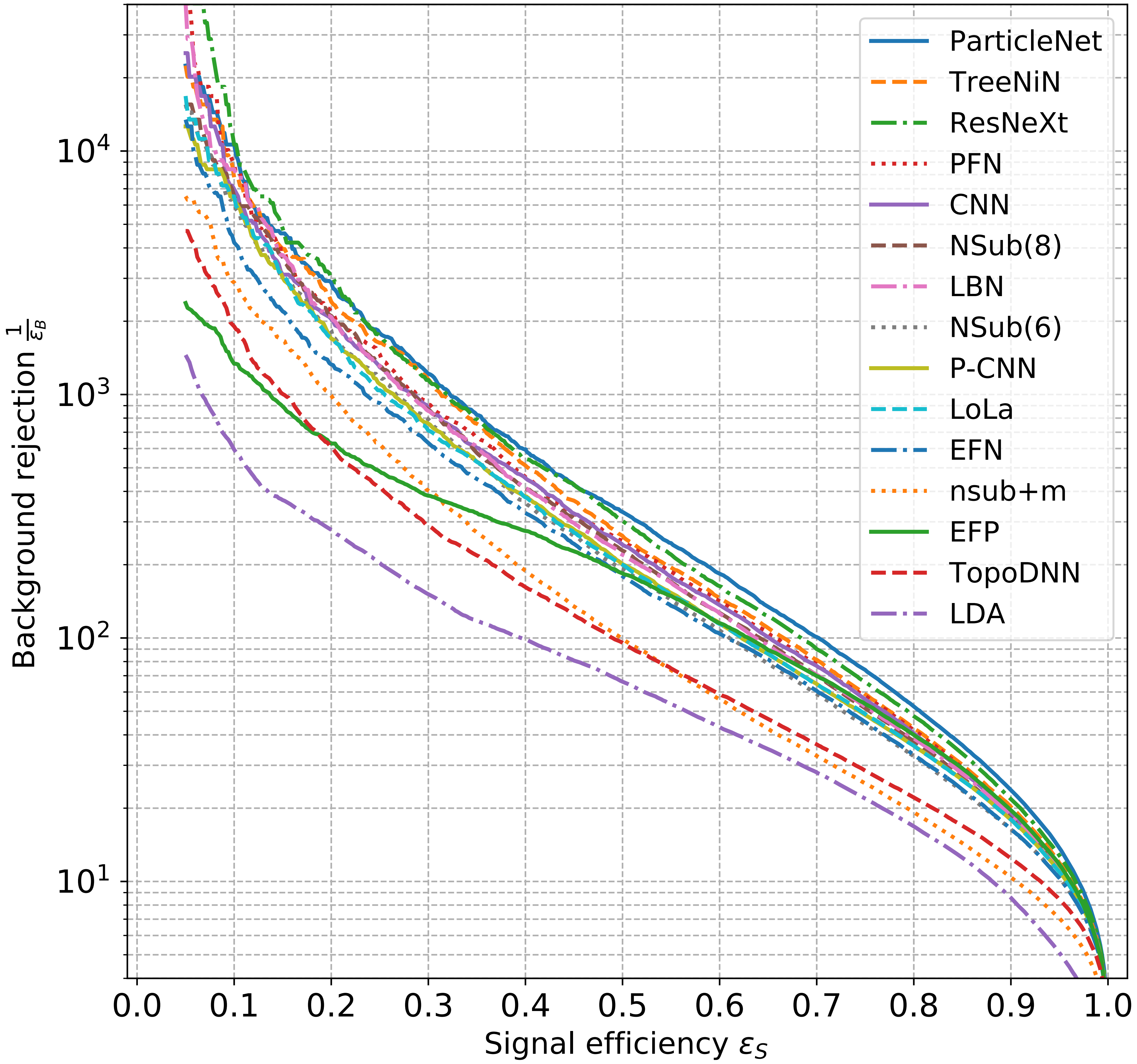
From 2015-2020 we saw rapid progress in using deep learning for

- Particle ID & jet / flavor tagging
- Reconstruction
  - Tracking, vertexing
  - Particle flow, pileup suppression
- Fast Simulation
- Anomaly Detection
- Unfolding
- Fast ML for Trigger

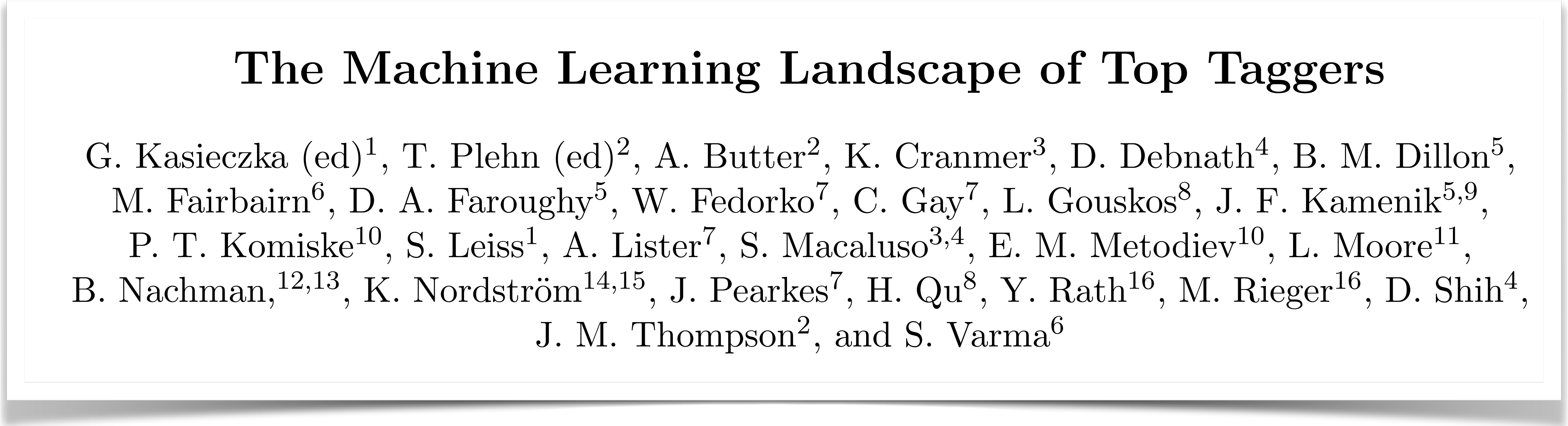
Now seeing a transition from early R&D prototypes to production

The Machine Learning Landscape of Top Taggers

G. Kasieczka (ed)[1], T. Plehn (ed)[2], A. Butter[2], K. Cranmer[3], D. Debnath[4], B. M. Dillon[5], M. Fairbairn[6], D. A. Faroughy[5], W. Fedorko[7], C. Gay[7], L. Gouskos[8], J. F. Kamenik[5,9], P. T. Komiske[10], S. Leiss[1], A. Lister[7], S. Macaluso[3,4], E. M. Metodiev[10], L. Moore[11], B. Nachman,[12,13], K. Nordström[14,15], J. Pearkes[7], H. Qu[8], Y. Rath[16], M. Rieger[16], D. Shih[4], J. M. Thompson[2], and S. Varma[6]

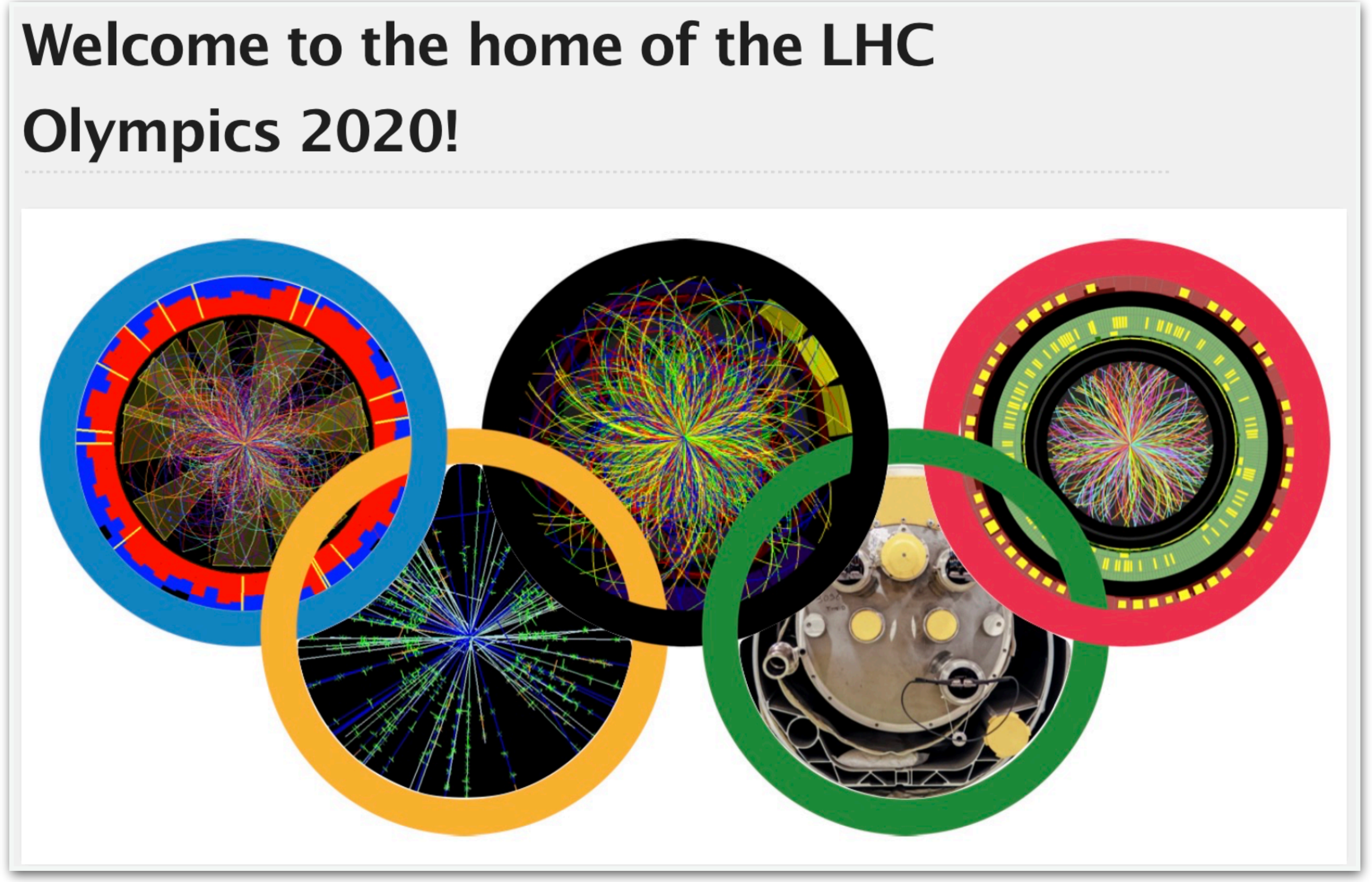


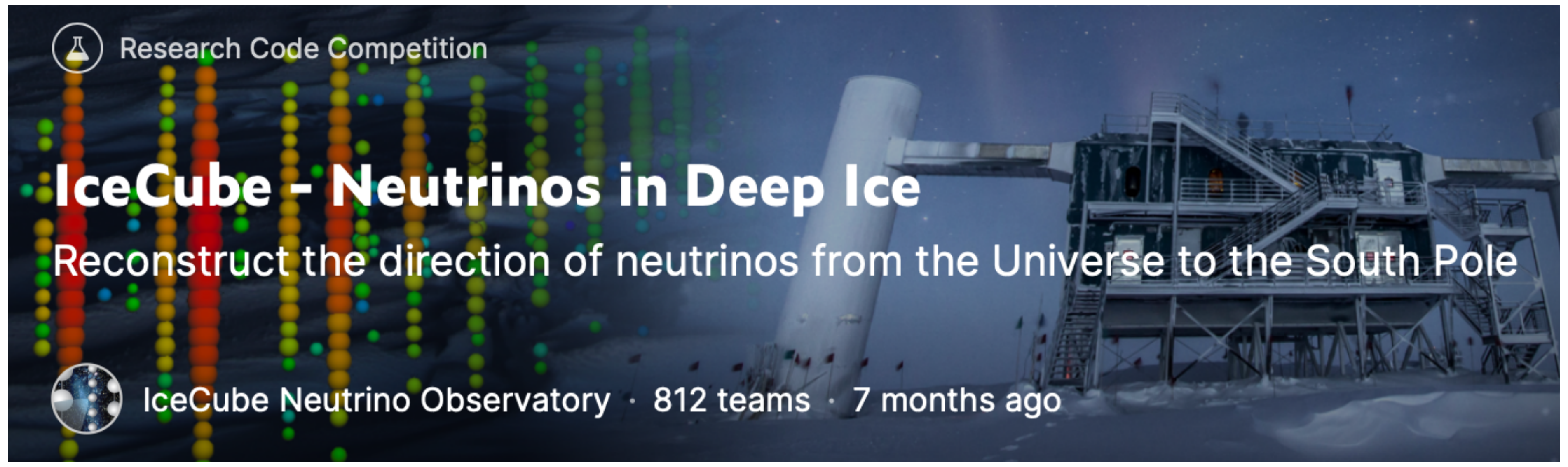
Challenges



The Machine Learning Landscape of Top Taggers

G. Kasieczka (ed)[1], T. Plehn (ed)[2], A. Butter[2], K. Cranmer[3], D. Debnath[4], B. M. Dillon[5], M. Fairbairn[6], D. A. Faroughy[5], W. Fedorko[7], C. Gay[7], L. Gouskos[8], J. F. Kamenik[5,9], P. T. Komiske[10], S. Leiss[1], A. Lister[7], S. Macaluso[3,4], E. M. Metodiev[10], L. Moore[11], B. Nachman,[12,13], K. Nordström[14,15], J. Pearkes[7], H. Qu[8], Y. Rath[16], M. Rieger[16], D. Shih[4], J. M. Thompson[2], and S. Varma[6]





Data Science at the Singularity

by David Donoho

Published on Jan 29, 2024

https://doi.org/10.1162/99608f92.b91339ef

- In the last decade, **frictionless services** became available thanks to the modern information ecosystem
- Those frictionless services were applied by scientists and technologists to **data sharing, code sharing, and challenges**
- Some communities of researchers started **frictionlessly sharing research artifacts** — code, data, results — and building on each others' work.
- Involved research communities are **progressing much faster**.

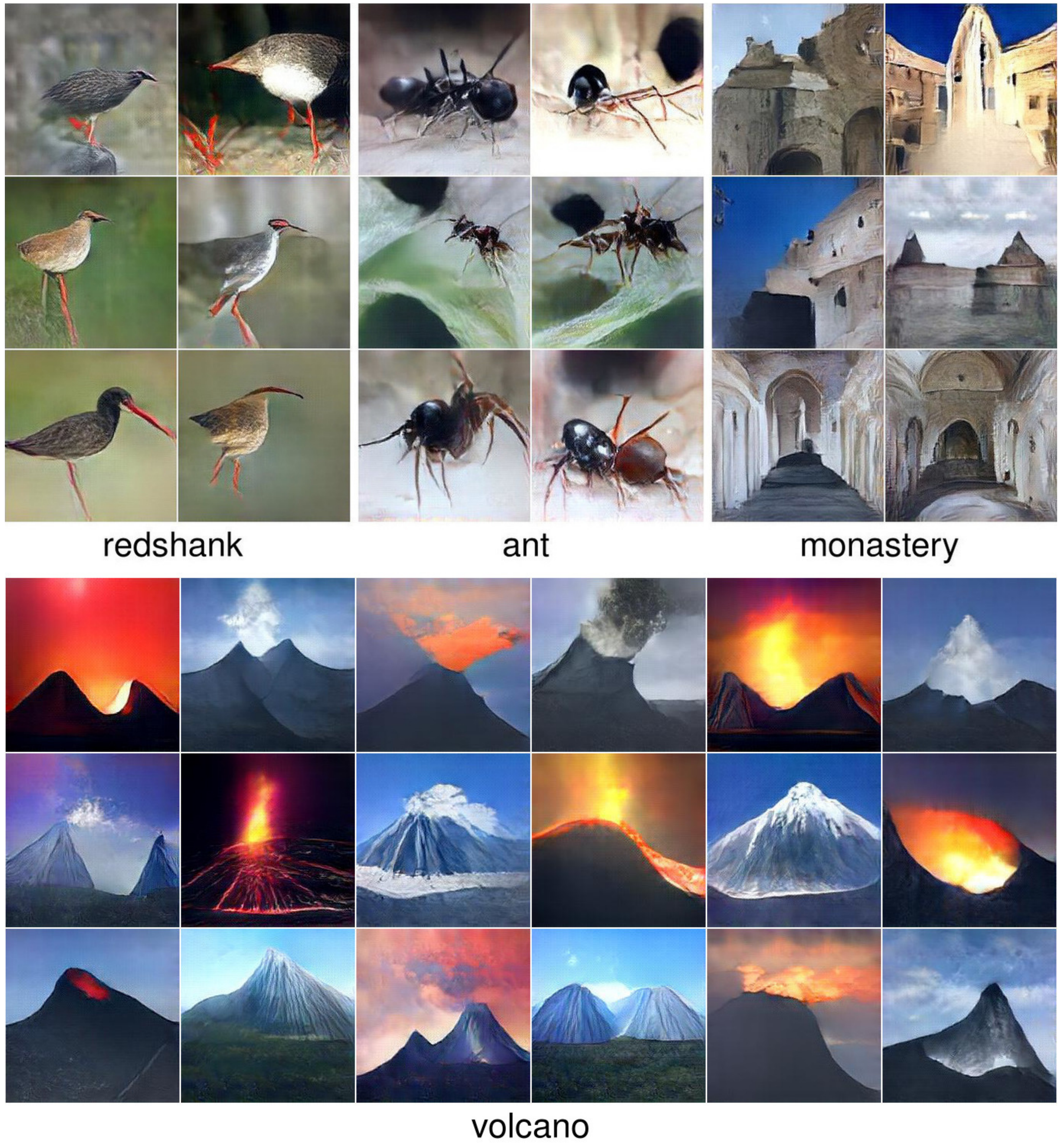
AI is one of those communities where people are working this way.
There is a **singularity**, but it is not AI.

The Evolution of Deep Learning

Predictive Models → Generative Models

Supervised learning → unsupervised learning

2016: Generative Model for Images



2018: Generative Model for Images

How an A.I. 'Cat-and-Mouse Game' Generates Believable Fake Photos

By CADE METZ and KEITH COLLINS JAN. 2, 2018

This one is computer-generated

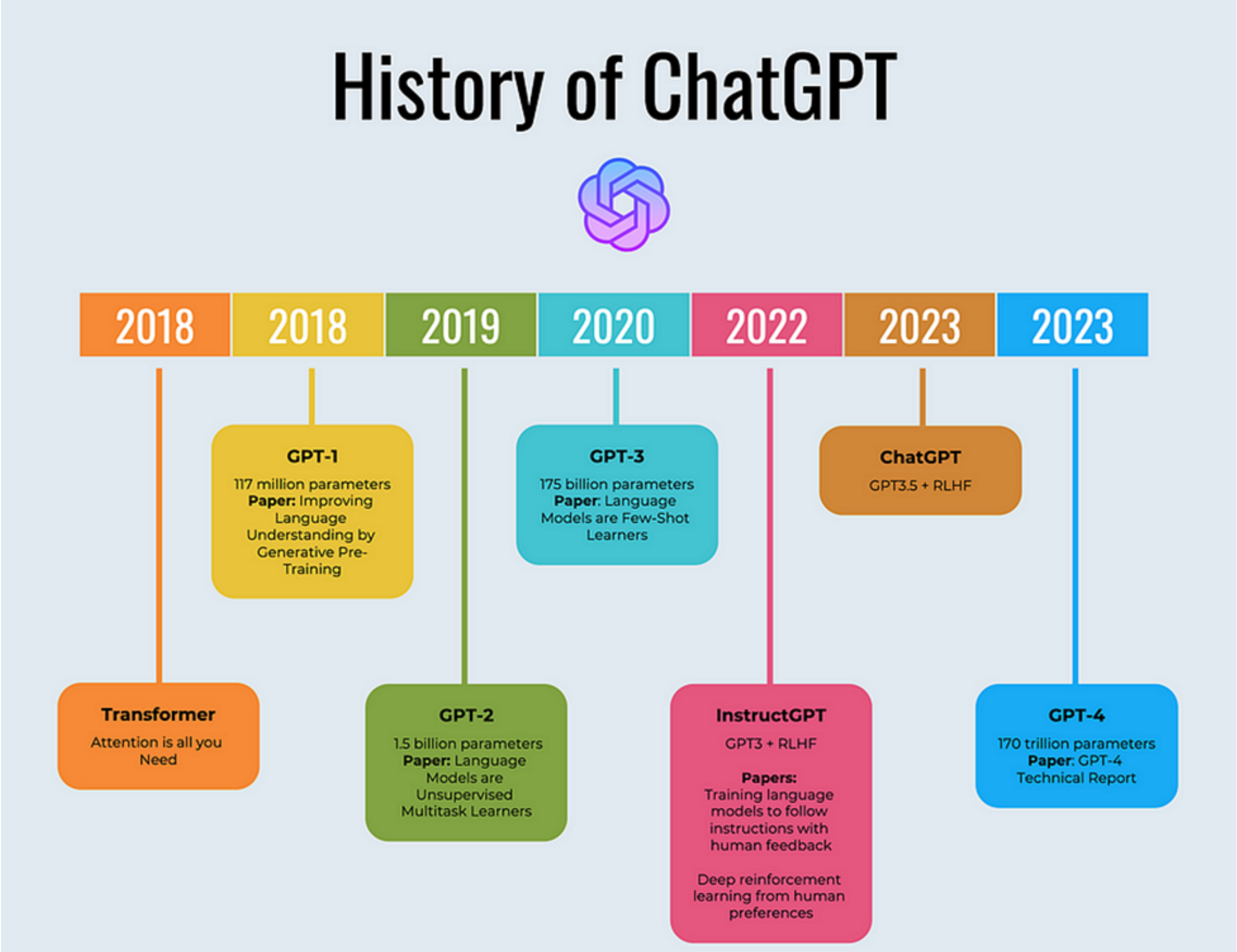
This one is also computer-generated

2018: Large Language Models

In 2018, a new approach to modeling language was introduced:

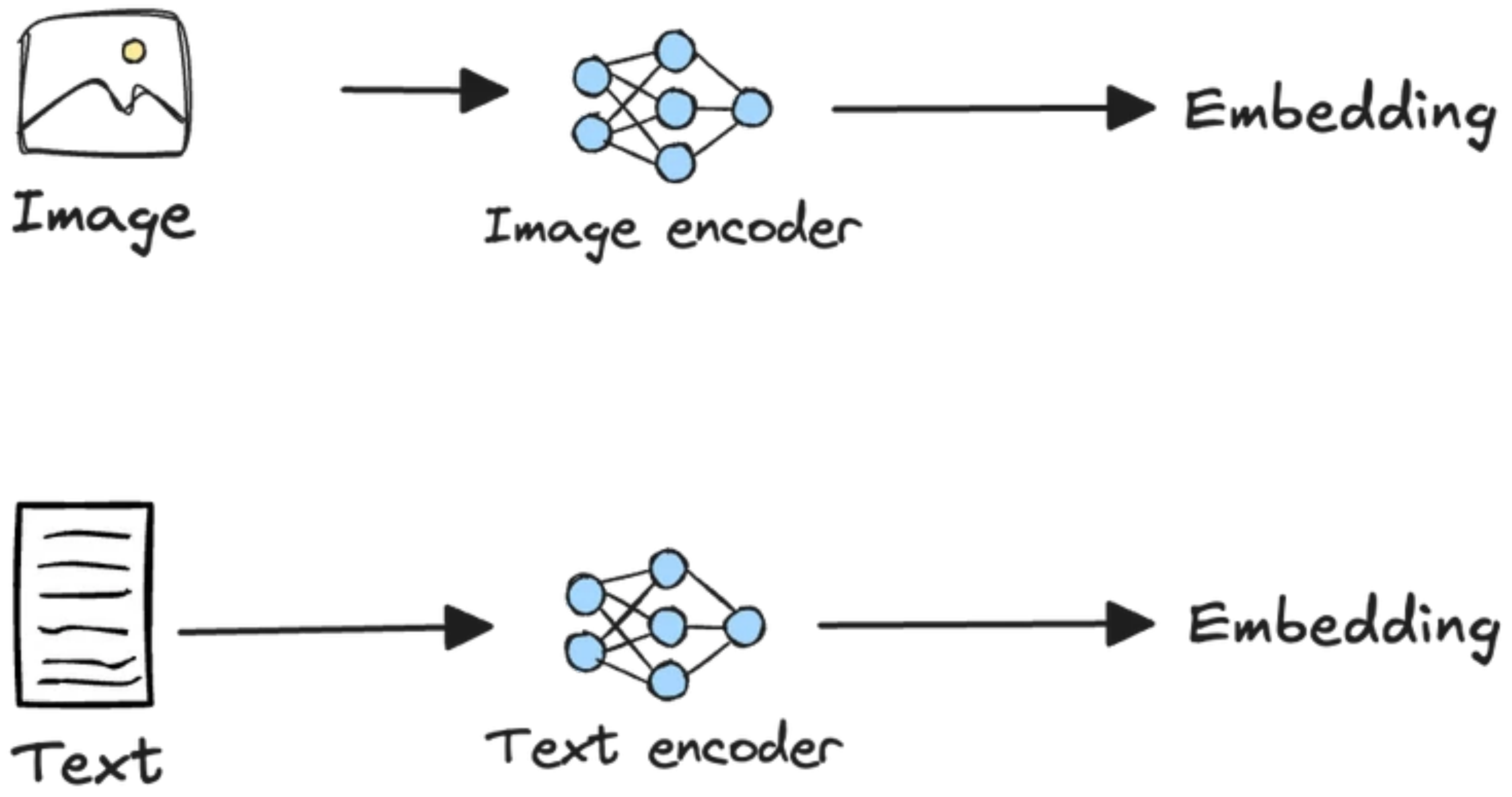
- The "transformer"
- Dramatic improvements
- Emergent capabilities
    - e.g. coding, "reasoning", ...

Transformers aren't specific to language, but they **provide bridge to domain knowledge in literature**



Multimodal modals

Major advance in creating models that can consume **multiple data modalities** and represent that data in a **shared semantic embedding space**



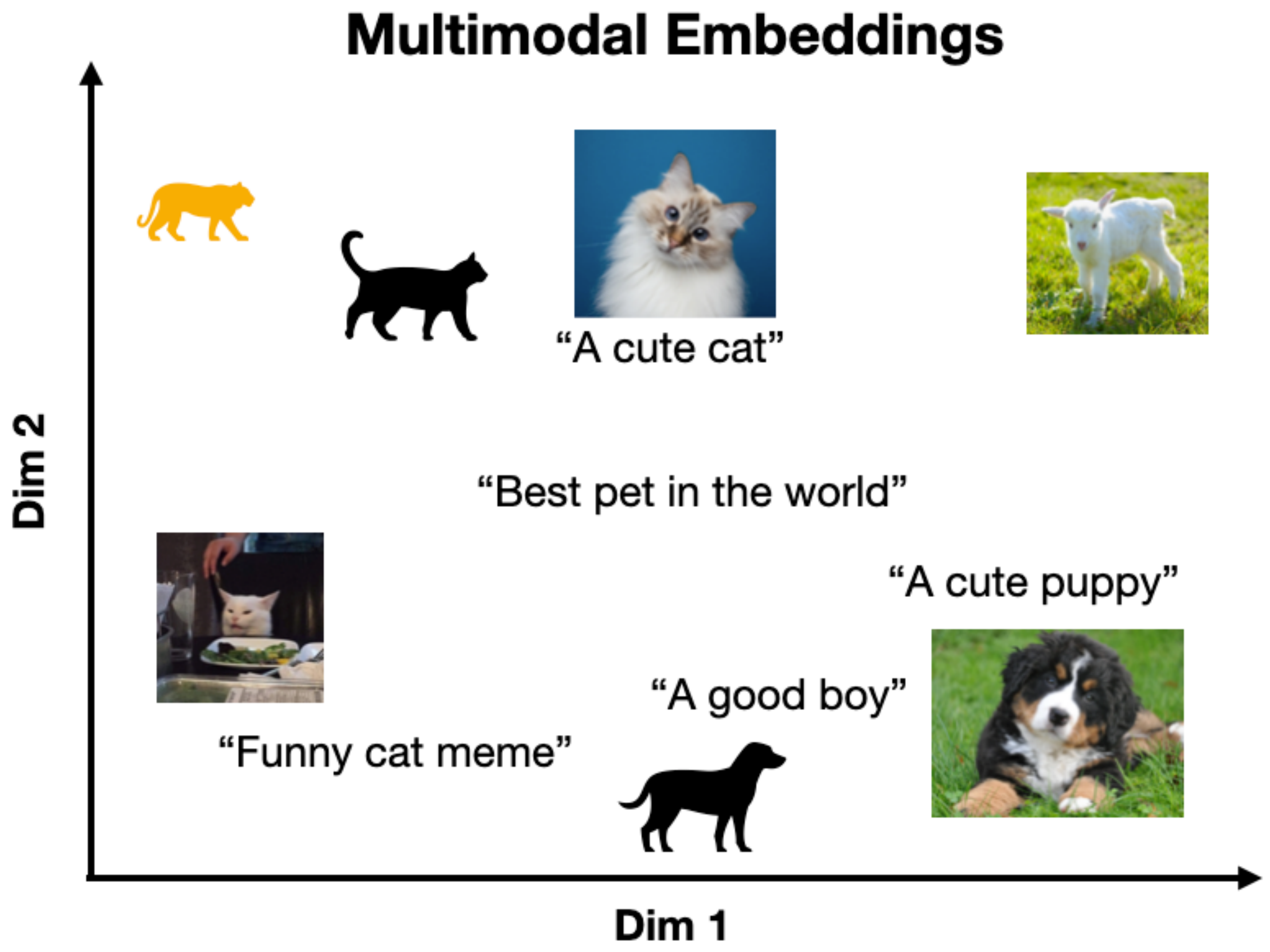
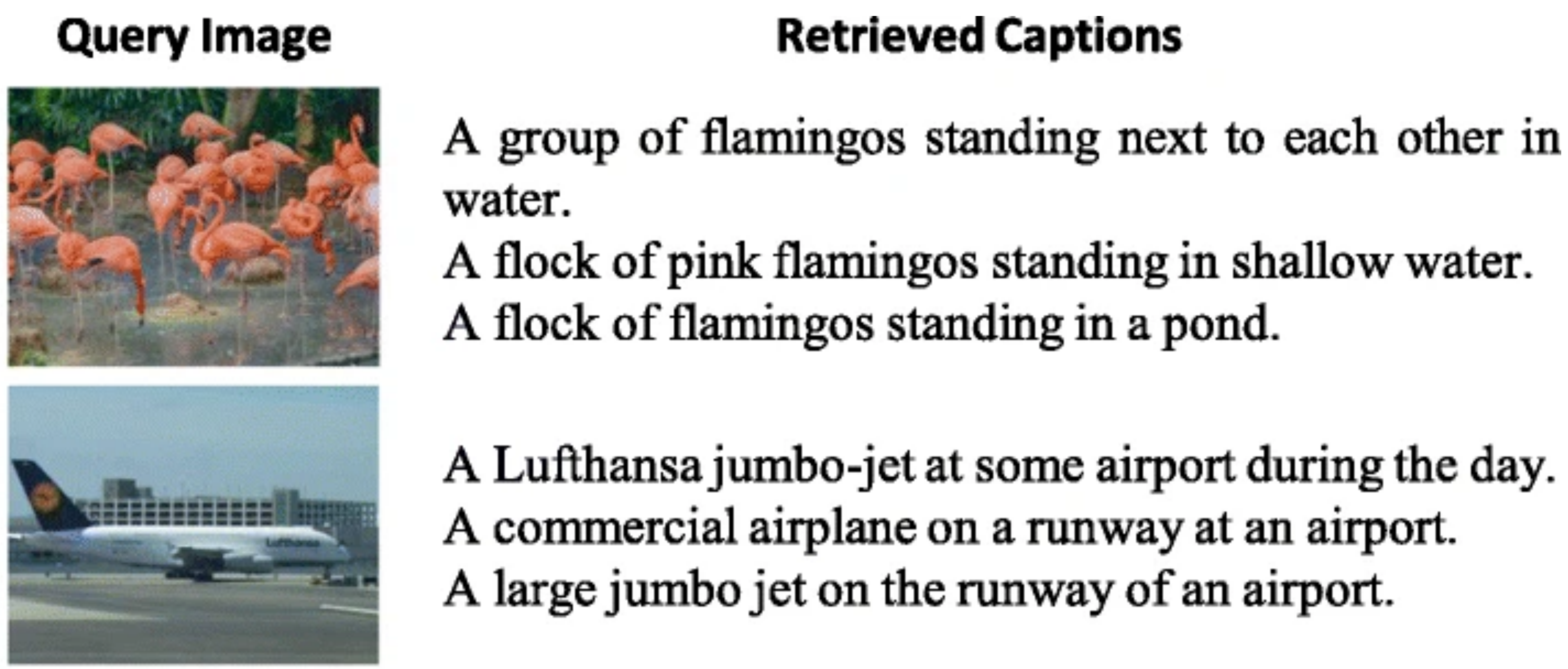


Figure source: 10.1007/s11042-020-09251-4

Multimodal modals

Major advance in creating models that can consume **multiple data modalities** and represent that data in a **shared semantic embedding space**

(a) Text-to-image retrieval



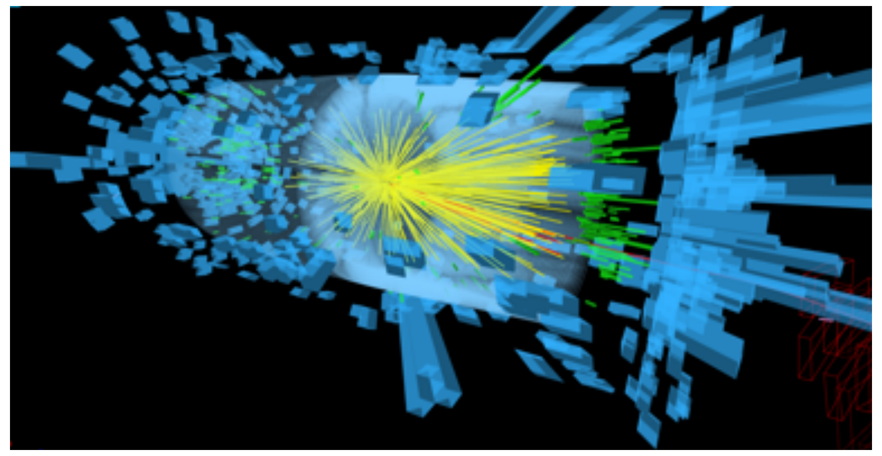
(b) Image-to-text retrieval

Figure source: 10.1007/s11042-020-09251-4

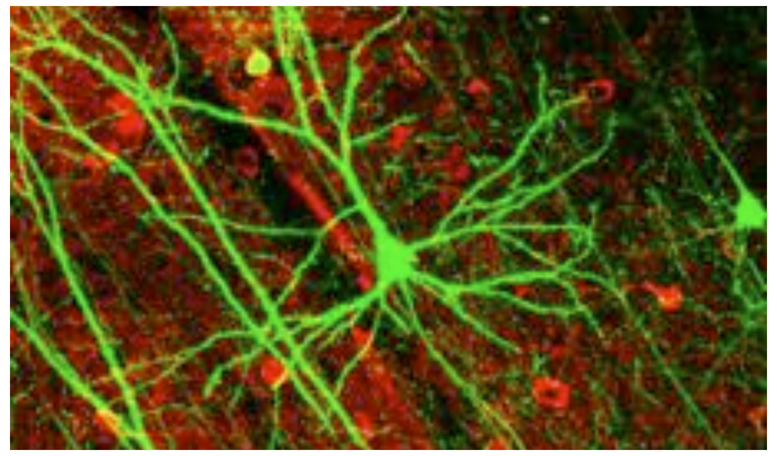
What are the dominant themes in AI for Science?

How does HEP fit in?

Simulators are the modern manifestation of theories



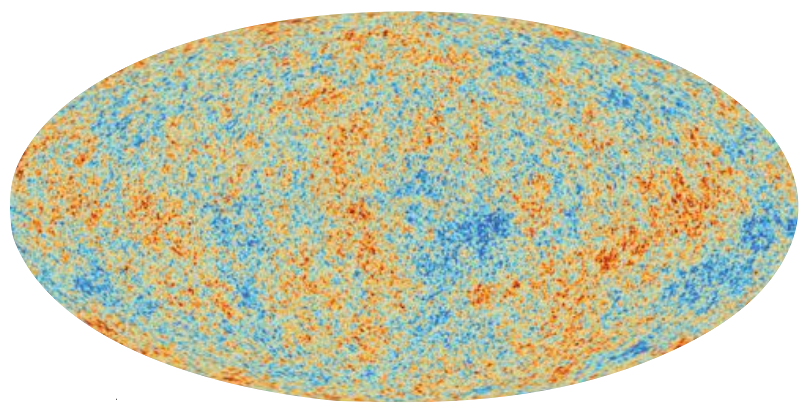
Particle colliders



Neuron activity

Epidemics

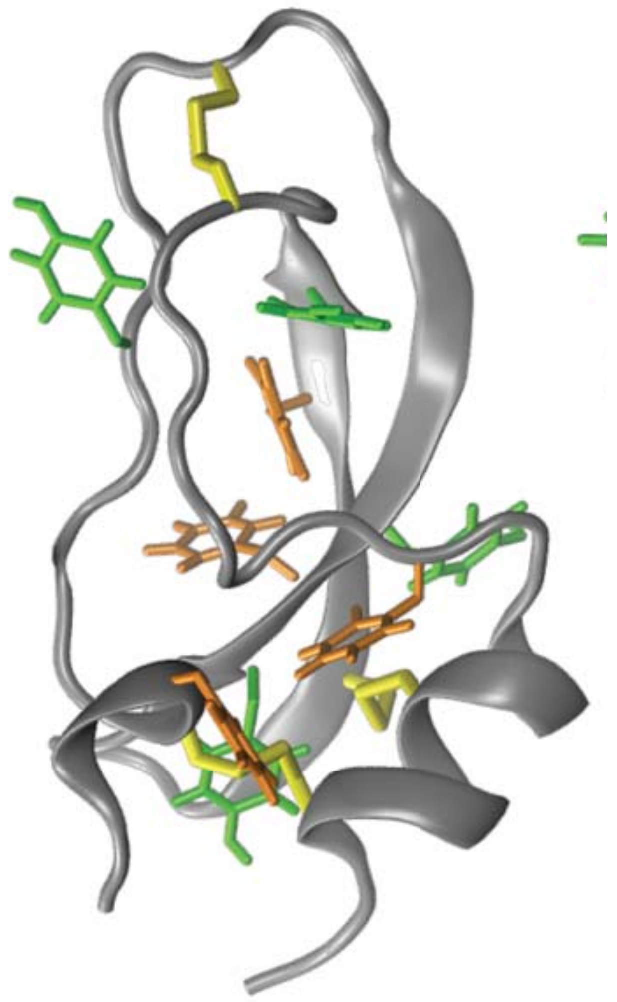
Gravitational lensing

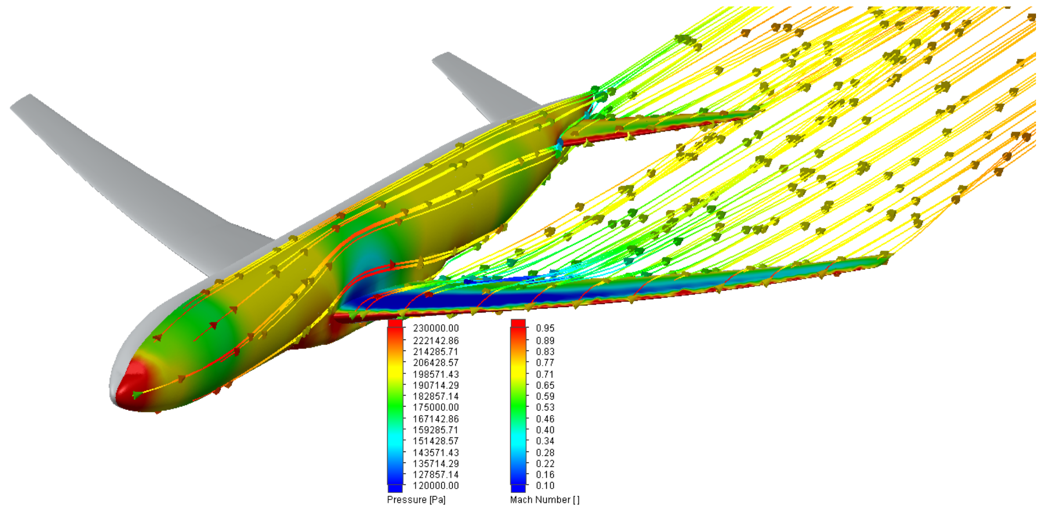


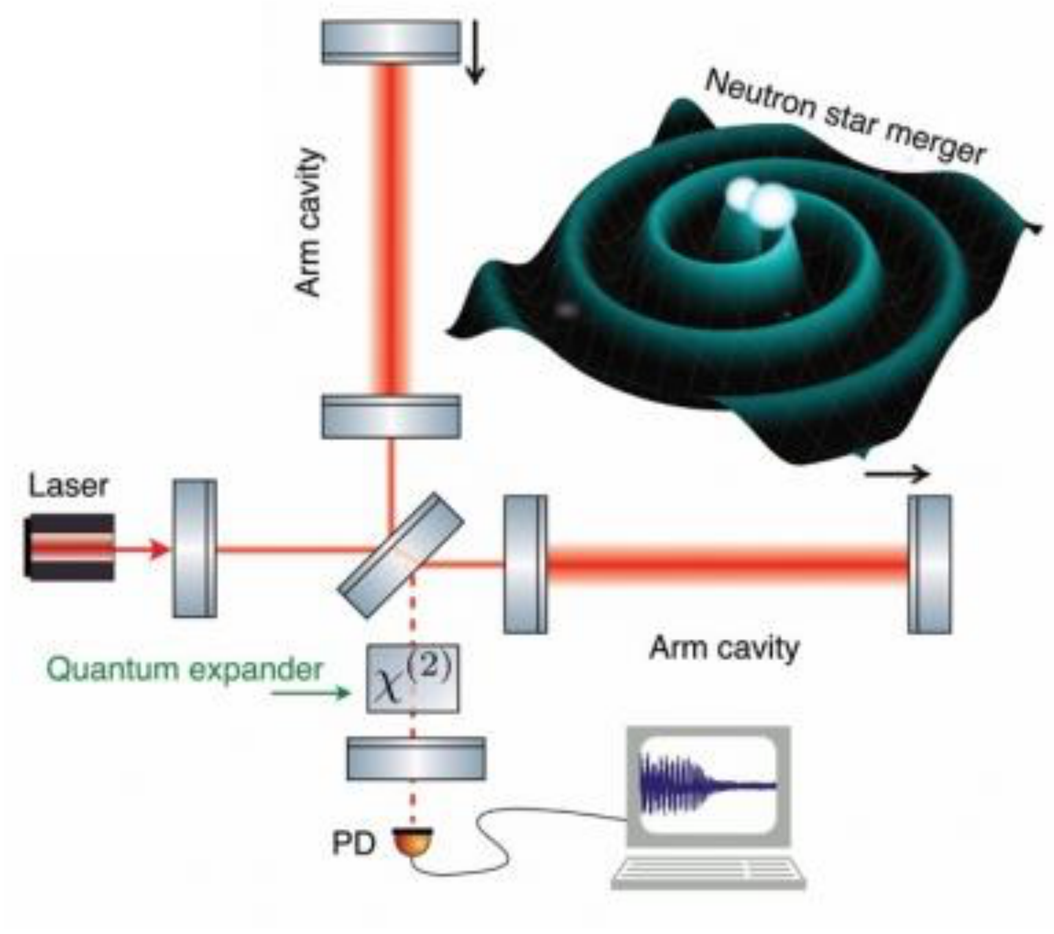
Evolution of the Universe

10^{-18} 10^{-15} 10^{-12} 10^{-9} 10^{-6} 10^{-3} 10^{0} 10^{3} 10^{6} 10^{9} 10^{12} 10^{15} 10^{18} 10^{21} 10^{24} 10^{27}

Length scale [m]



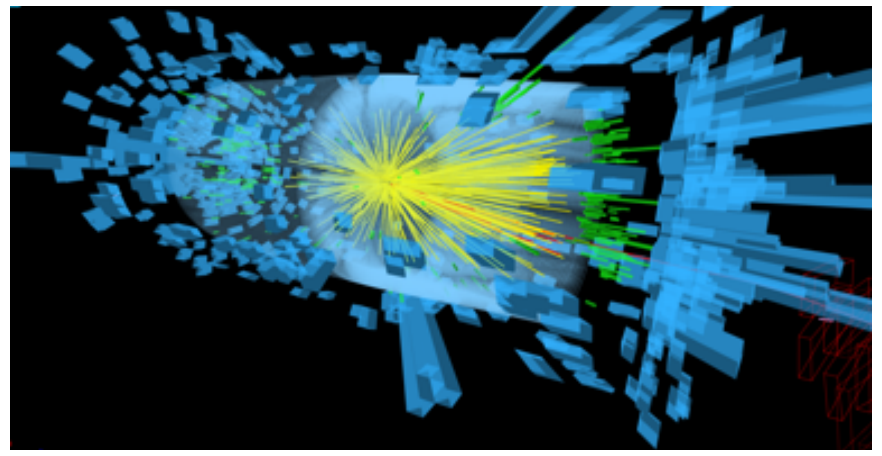




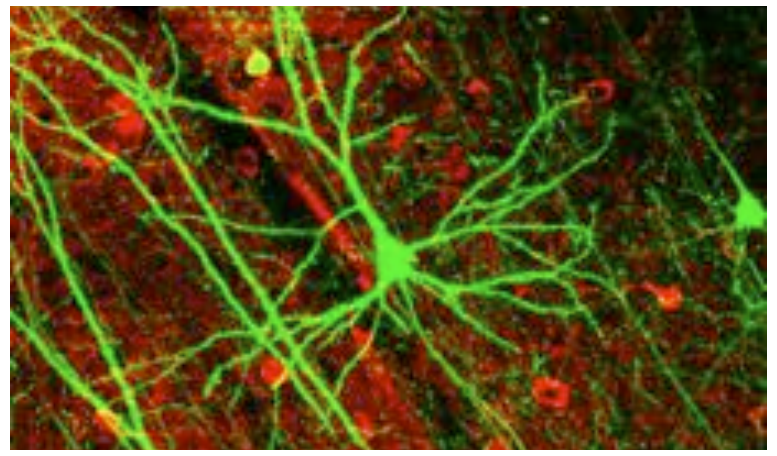
The forefront of scientific knowledge is often encapsulated in simulators

[Cranmer, Brehmer, Louppe PNAS (2020), arXiv:1911.01429]

Simulators are the modern manifestation of theories



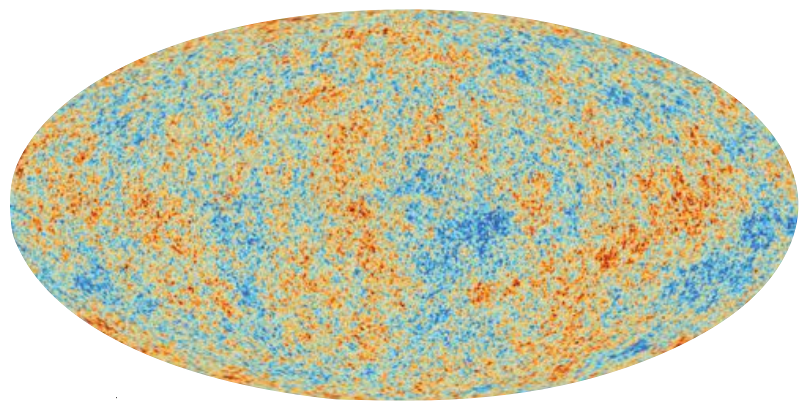
Particle colliders



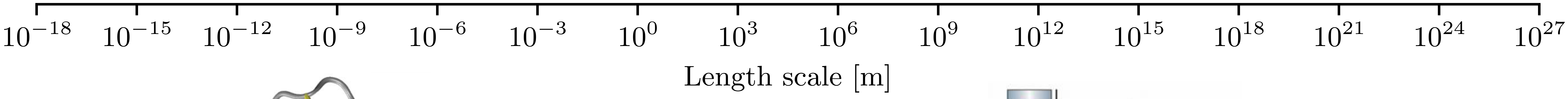
Neuron activity

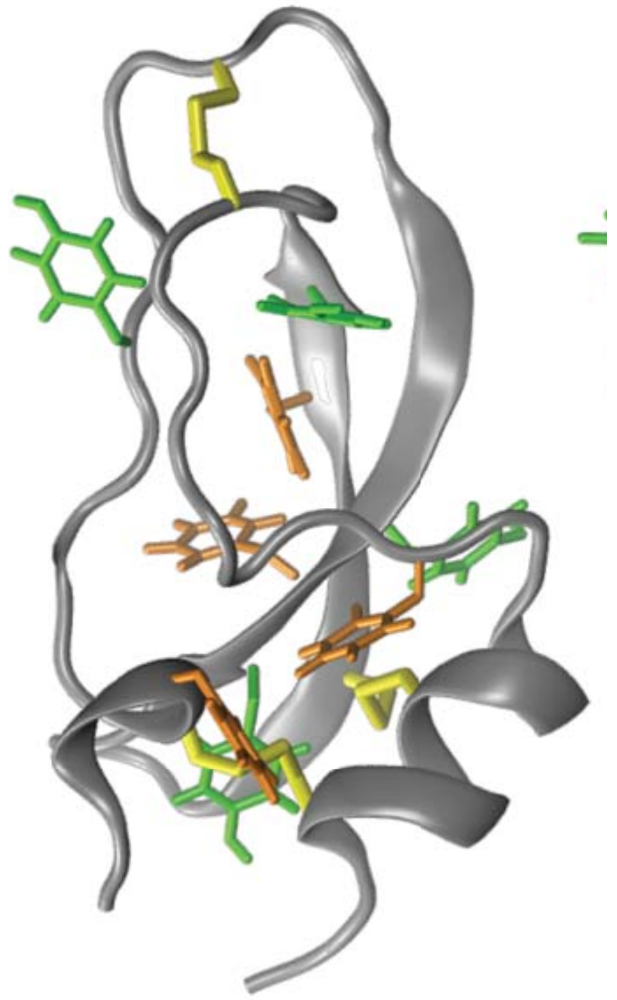
Epidemics

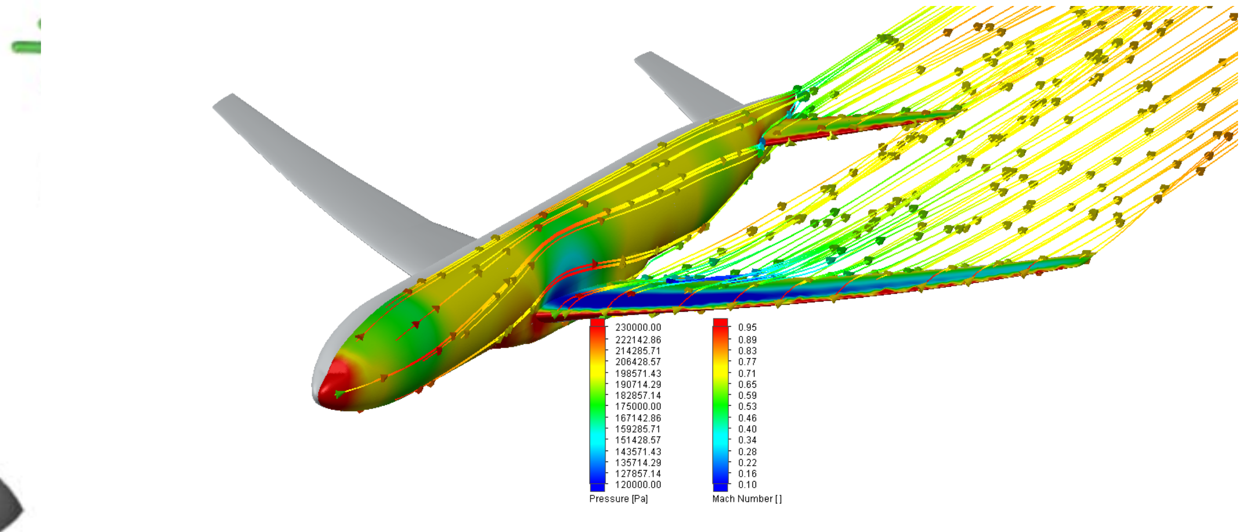
Gravitational lensing

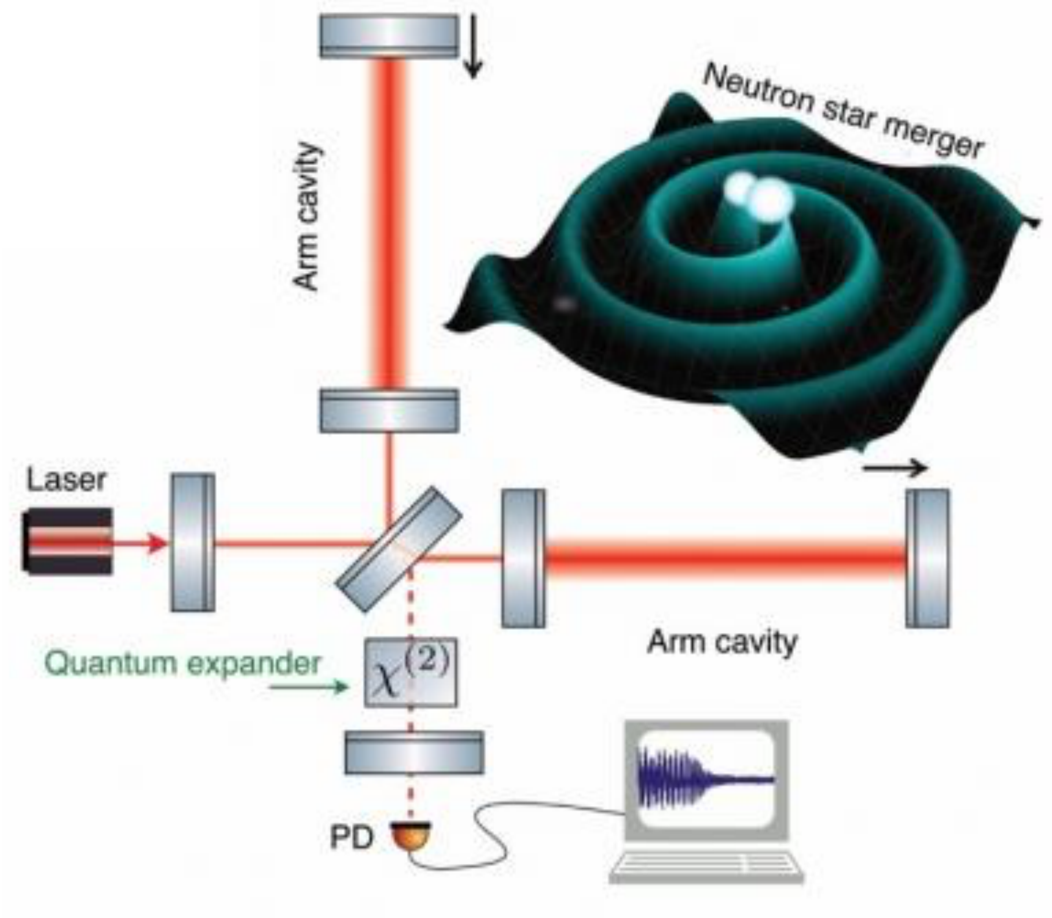


Evolution of the Universe







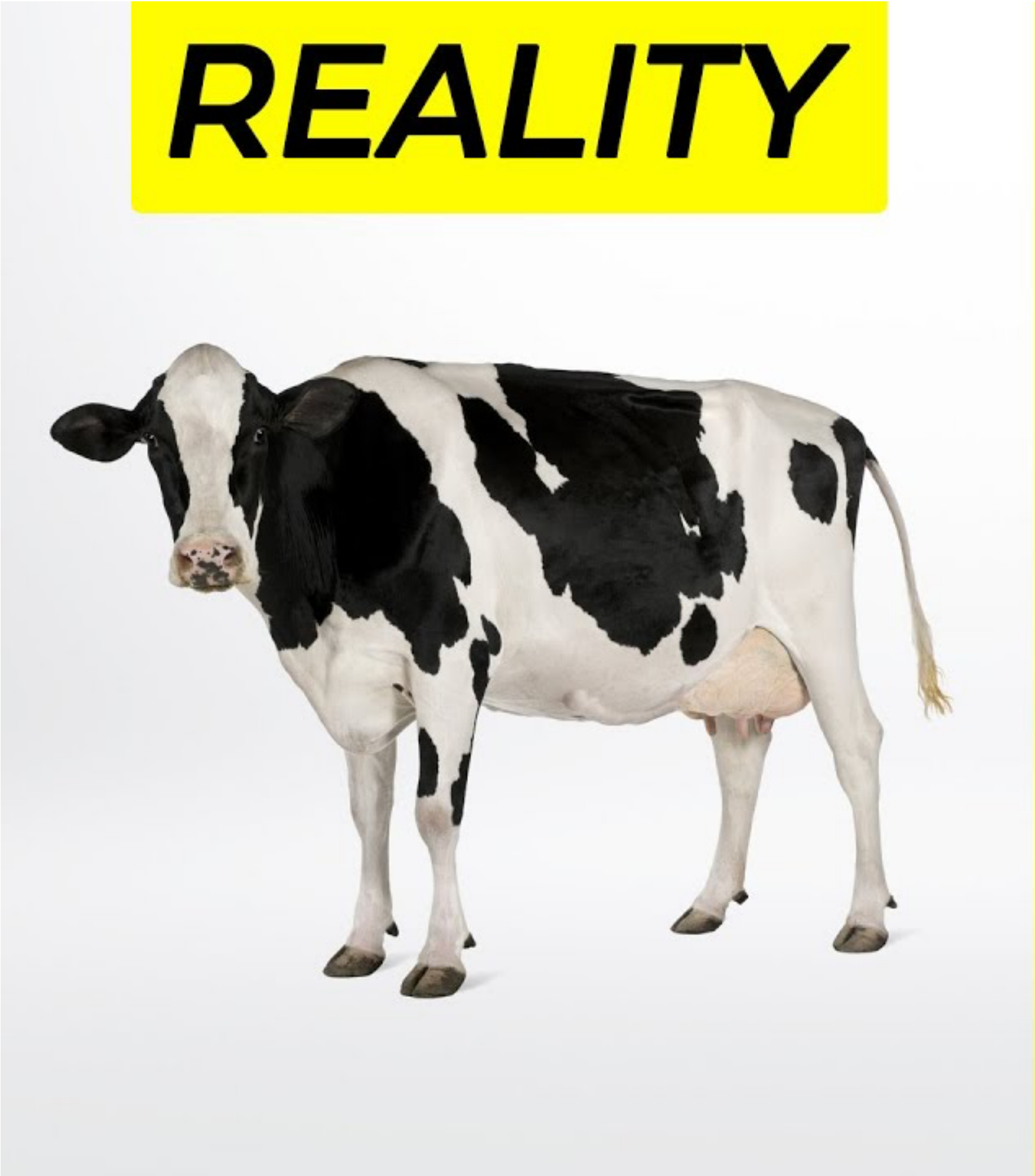


Unfortunately, simulators are poorly suited for many downstream tasks, e.g. statistical inference, experimental design, decision making, …

[Cranmer, Brehmer, Louppe PNAS (2020), arXiv:1911.01429]

"The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and **the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble.**"

–PAUL DIRAC



Cow

Spherical Cow

REALITY

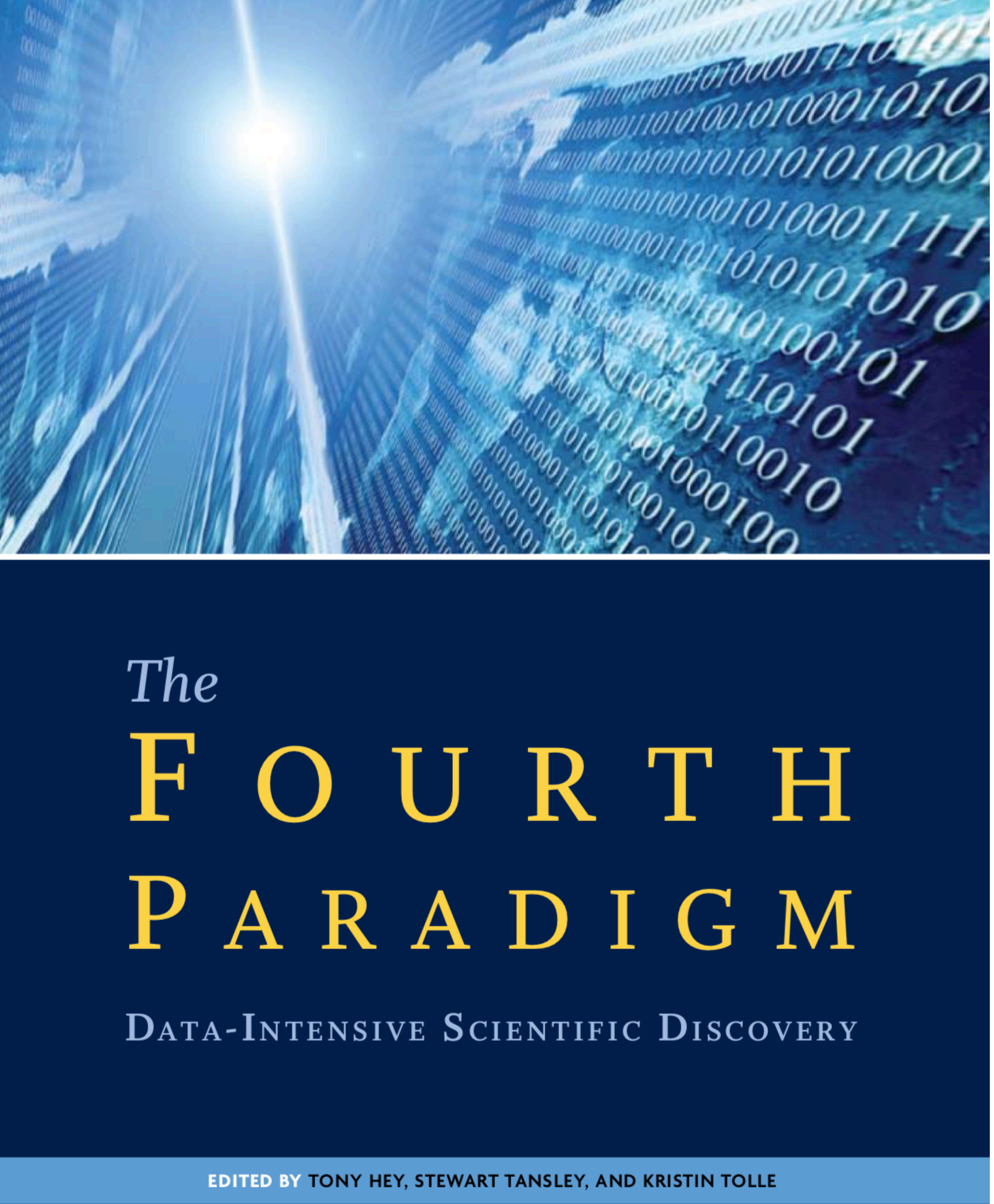
Cow

PHYSICS

Spherical Cow

AI+SCIENCE

CowFormer



Scientific Paradigms:

1. Empirical
2. Theoretical
3. Computational
4. Data-Driven / Data-Intensive Discovery
5. **AI/ML + Simulation + Data**

AI4Science to empower the fifth paradigm of scientific discovery

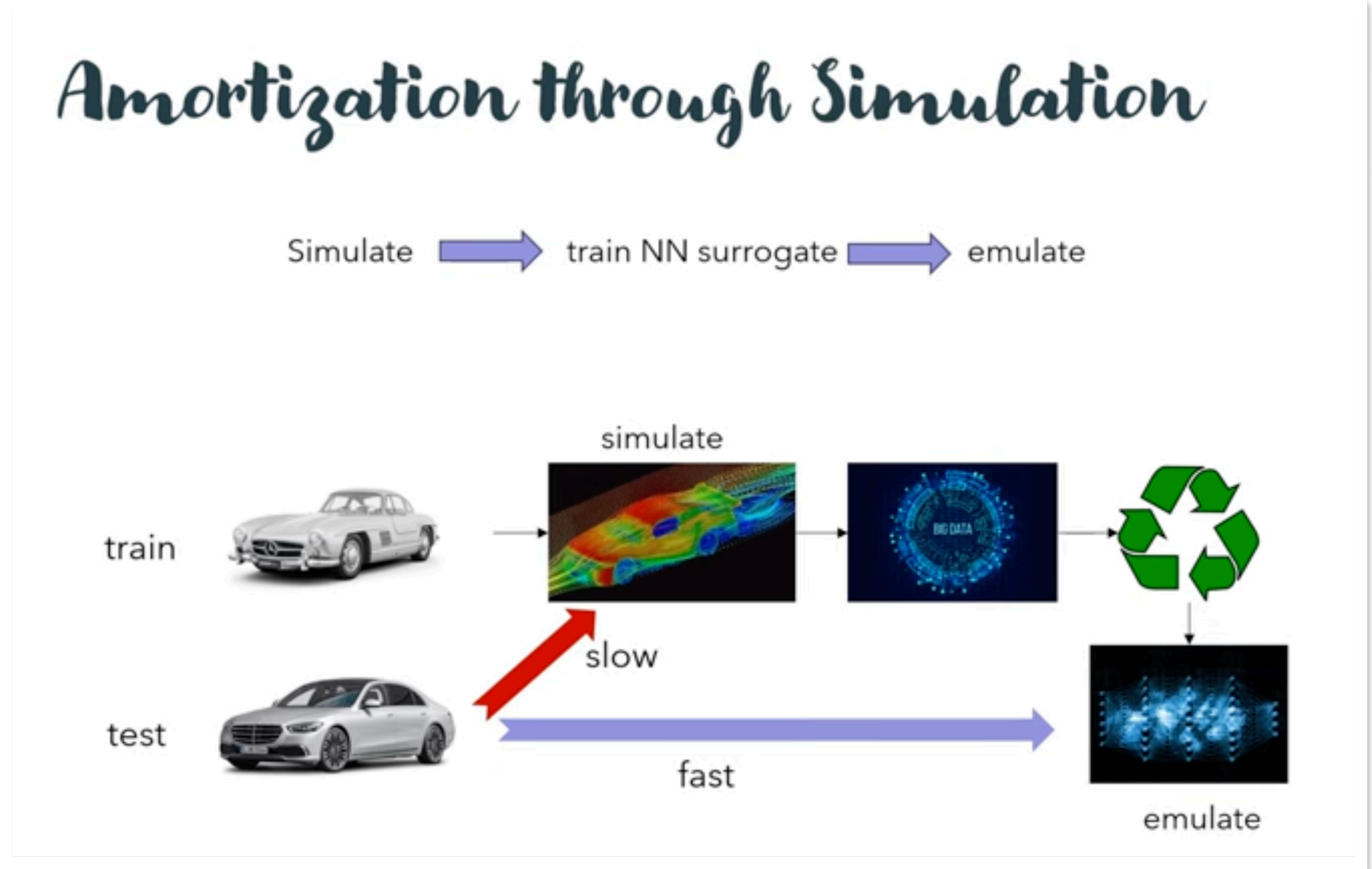
Published July 7, 2022

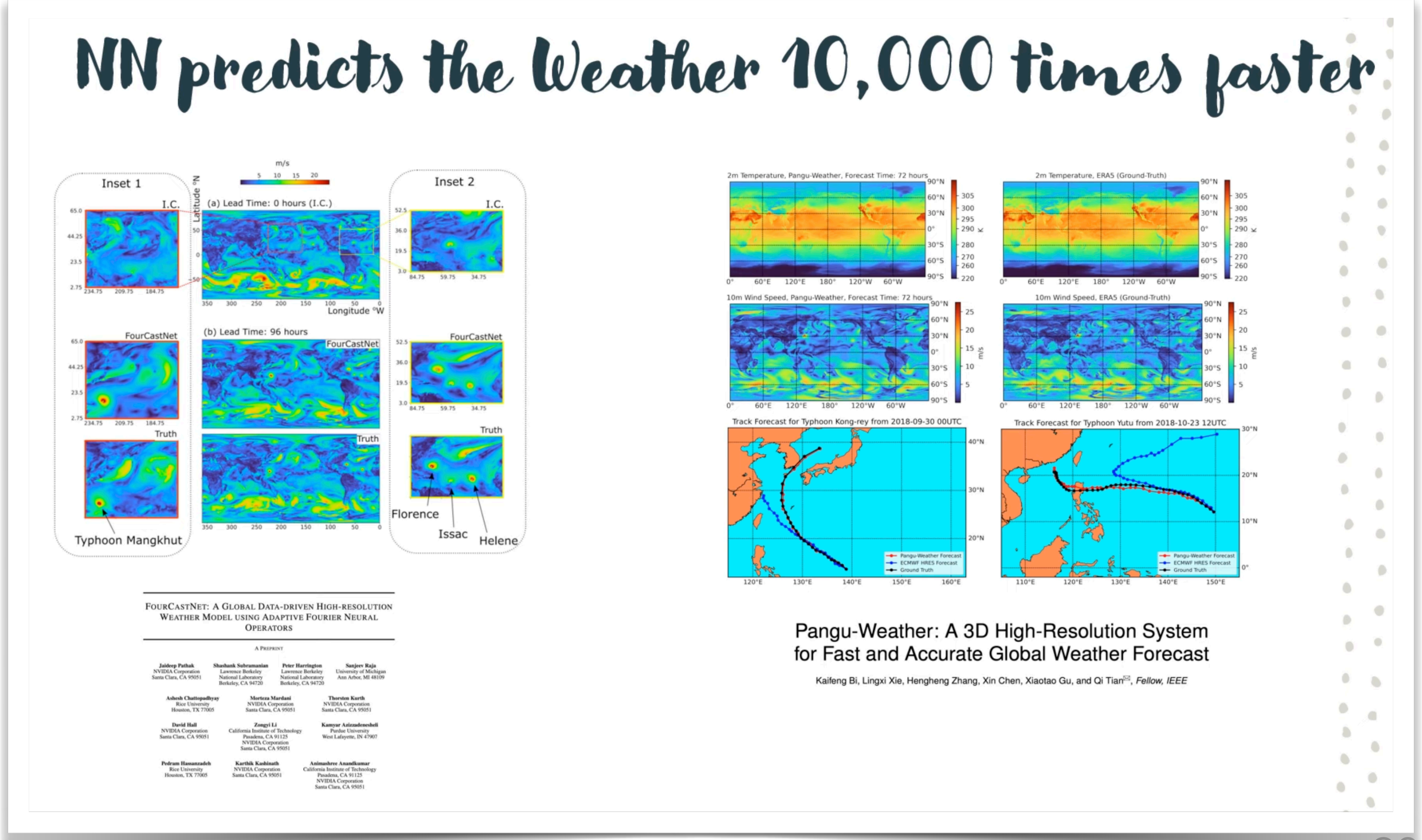
By Christopher Bishop, Technical Fellow and Director, Microsoft Research AI4Science

Share this page

https://www.microsoft.com/en-us/research/blog/ai4science-to-empower-the-fifth-paradigm-of-scientific-discovery/

Simulation & Emulators



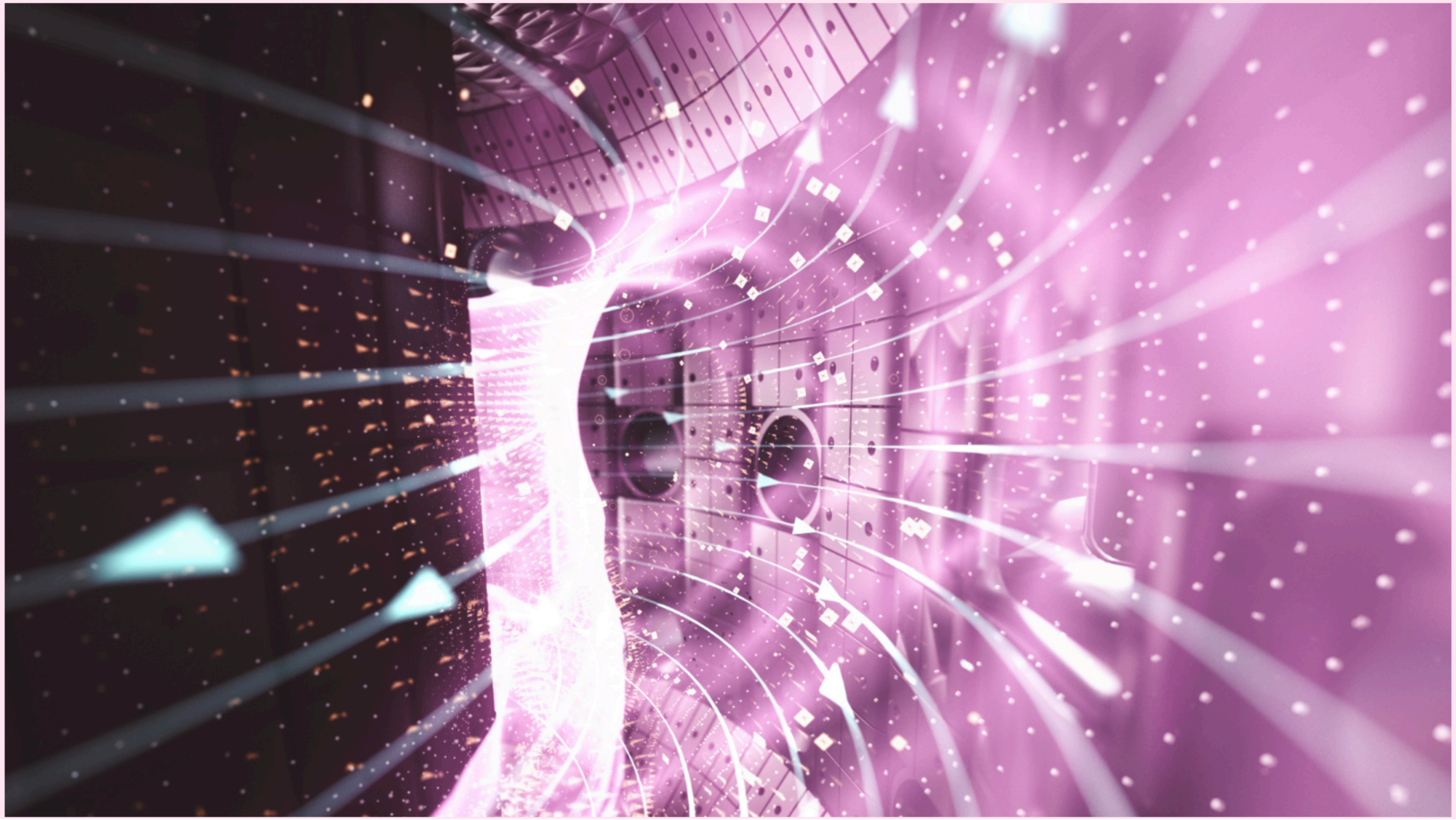


In fusion energy

Research

Accelerating fusion science through learned plasma control

February 16, 2022

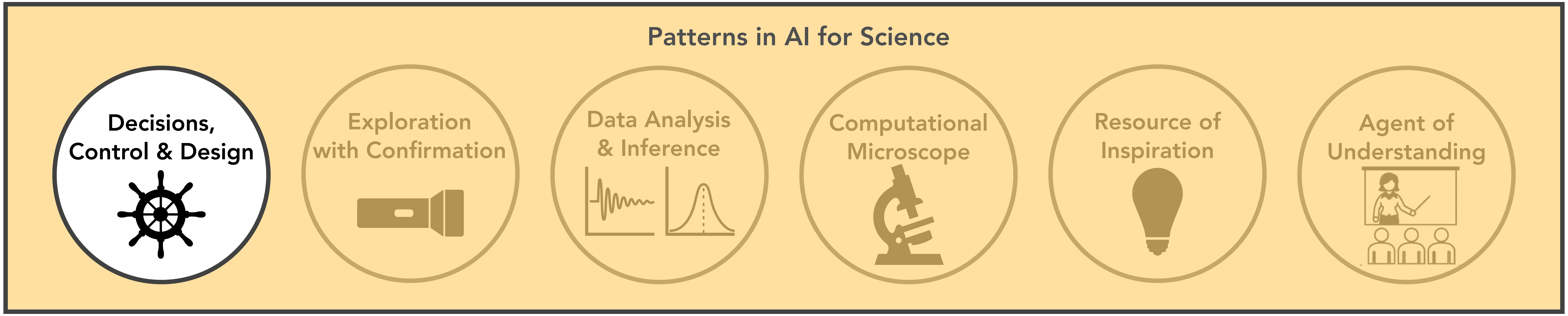


Successfully controlling the nuclear fusion plasma in a tokamak with deep reinforcement learning

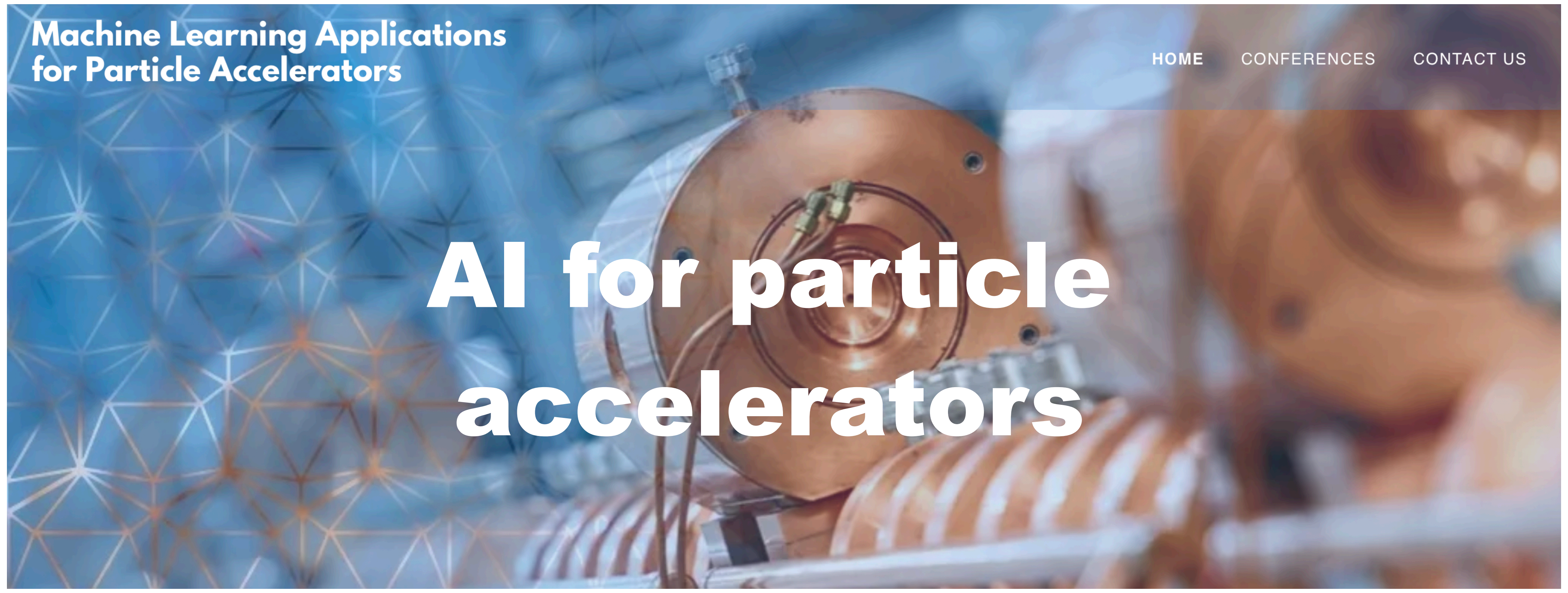
AI/ML is providing dramatically enhanced capabilities

Fast AI/ML emulators of classical numerical simulations enable these systems to be used for **decisions, control, & design** where it was previously infeasible

- Numerical weather prediction / fusion / …
- But also in HEP

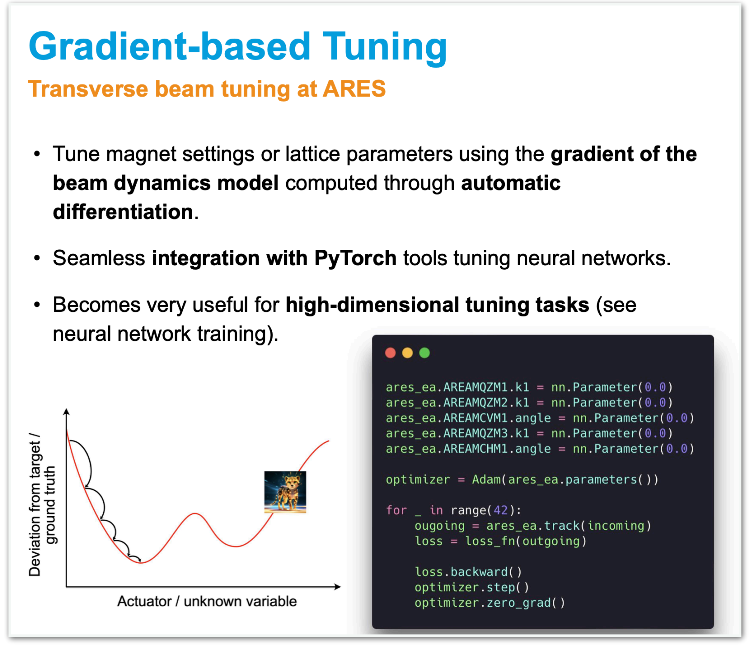


Verena Kain's talk @ EuCAIF

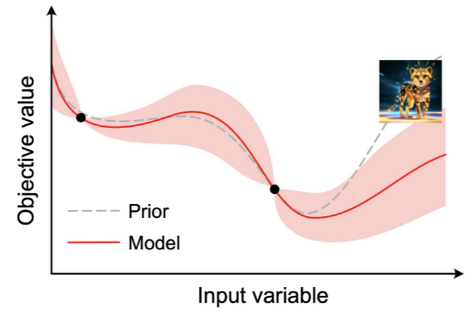


Differentiable simulation codes

Optimisation algorithms work best and are most sample-efficient with gradient information of the objective function.



- A physics-informed prior can help **improve the performance of BO** by preventing over-exploitation.
- Cheetah's differentiability allows **efficient acquisition function optimisation** using gradient descent methods in modern BO packages like BoTorch.
- Has well-defined behaviour and **does not need data** to train like neural network priors.
- Can be used in **combination with gradient-based system identification** to overcome model inaccuracies.



RL4AA - workshop

Pushing the frontiers of RL for accelerators → autonomous accelerators.

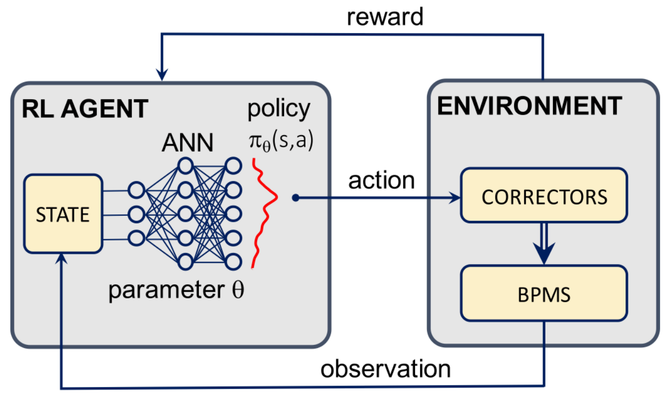
RL4AA Collaboration

The Reinforcement Learning for Autonomous Accelerators international collaboration aims to consolidate the existing knowledge in the community, exchange experience and ideas, and work together towards accelerator-specific solutions using the latest advances in RL

www.youtube.com

Reinforcement Learning (RL)

Learn dynamics (once and for all) through trial-and-error, no exploration after training!



RL setup for trajectory steering

Next generation accelerators to be built for RL:
→ fast executing (accurate) simulation / digital twin for training
→ instrumentation designed with control algorithm

RL elegant (if not ideal) solution, but **online training** often not possible!

- Not sample-efficient enough
- Safety constraints

→ **RL (like MPC) needs to be built into accelerator design.**

RL for accelerator operations at DESY

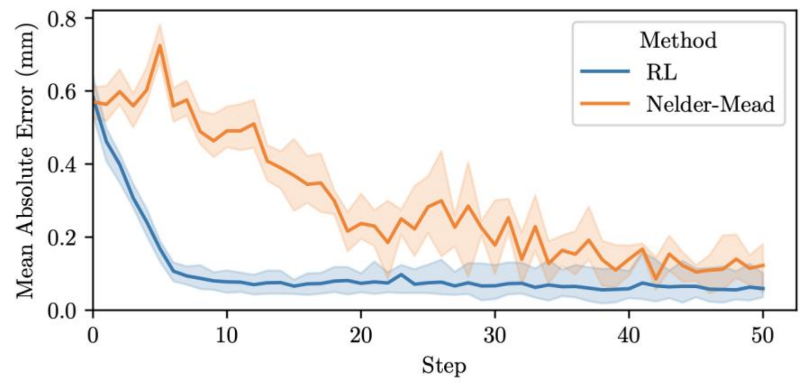
RL can tune 4x faster than human operators

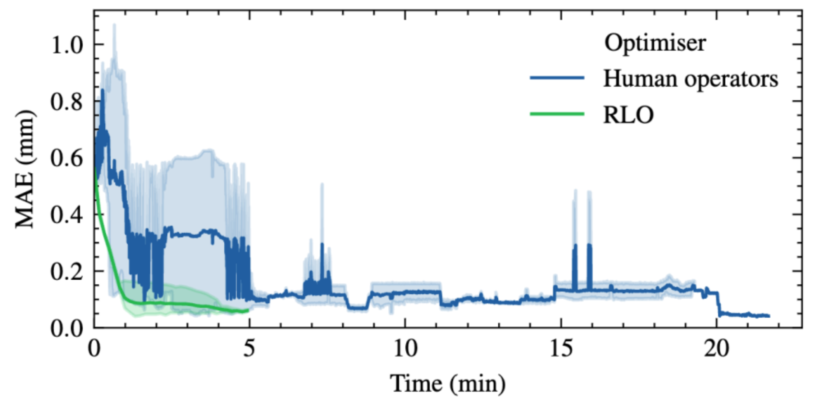
MATTER AND TECHNOLOGIES MT DESY.

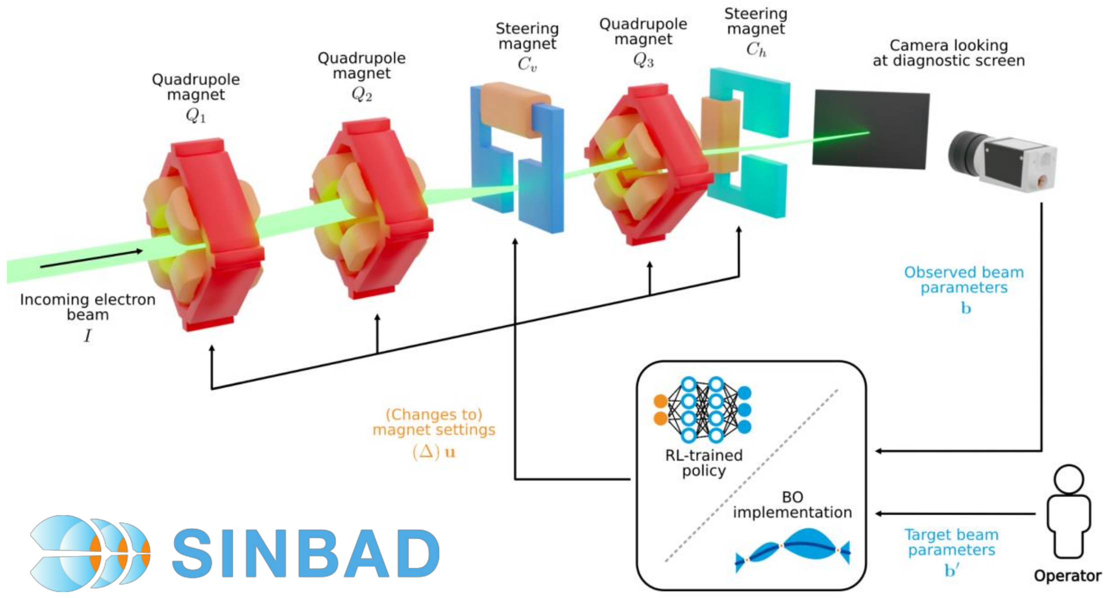
Reinforcement learning: From ARES Sinbad to the European XFEL

Reinforcement learning-trained optimization at ARES

- Deploy a RL-trained optimization algorithm trained purely in simulation to the **real-world** with **zero-shot learning** thanks to **domain randomization.**
- The trained policy **outperforms other optimization algorithms and expert human operators**.







J Kaiser, O Stein, A Eichler. **Learning-based Optimisation of Particle Accelerators Under Partial Observability Without Real-World Training**. In *International Conference on Machine Learning*, 2022.
J Kaiser, C Xu, A Eichler, et. al. **Reinforcement learning-trained optimisers and Bayesian optimisation for online particle accelerator tuning**. In *Scientific reports* 14 (1), 2024

Cheetah: Speeding up simulations by 10^8

MATTER AND TECHNOLOGIES MT DESY.

Linear beam dynamics simulation python package

Why we need it? Training of RL agents would require **3 years of beam time** on the real machine, **11 days with Ocelot, 1 hour with Cheetah.**

Main features in support of ML applications:

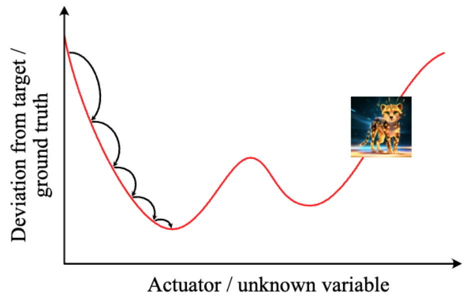
- **Ultra-fast** compute (at the cost of fidelity)
- **Differentiability**
- **GPU support**

TABLE I. Step computation times of simulation codes in milliseconds

Code	Comment	Laptop	HPC node
ASTRA	space charge	264 000.00	3 605 000.00
	no space charge	109 000.00	183 000.00
Parallel ASTRA	space charge	39 000.00	17 300.00
	no space charge	16 900.00	12 600.00
Ocelot	space charge	22 100.00	21 700.00
	no space charge	182.00	119.00
Bmad-X		40.50	74.30
Cheetah	`ParticleBeam`	0.79	0.72
	`ParticleBeam` + GPU	-	0.09
	`ParameterBeam`	0.02	0.04

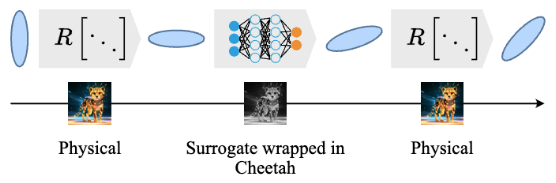
Making use of Cheetah's **differentiability**

- Bayesian optimization prior
- Gradient-based tuning / system identification



Making use of Cheetah's **speed**

- Reinforcement learning
- Integration of modular network surrogate



Cheetah in daily operation at LCLS:

SLAC NATIONAL ACCELERATOR LABORATORY

Now deployed to **daily** operations at **LCLS** for 6D-phase space reconstruction module.

Jan Kaiser, Chenran Xu, Annika Eichler and Andrea Santamaria Garcia. **Bridging the Gap Between Machine Learning and Particle Accelerator Physics with High-Speed, Differentiable Simulations**. In *Physical Review Accelerators and Beams*, 2024.

RL for accelerator operations at DESY

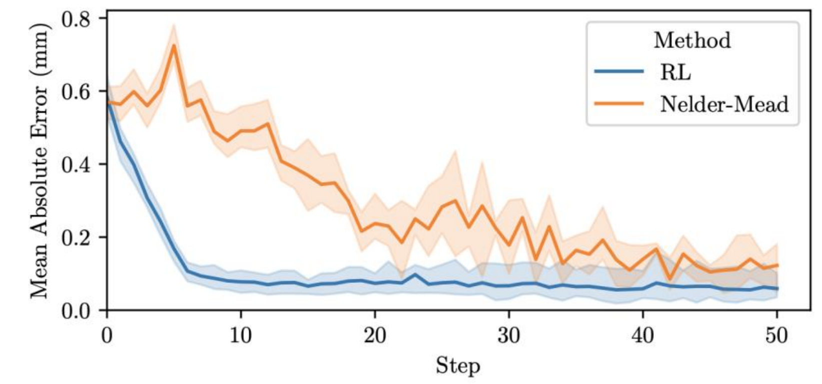
RL can tune 4x faster than human operators

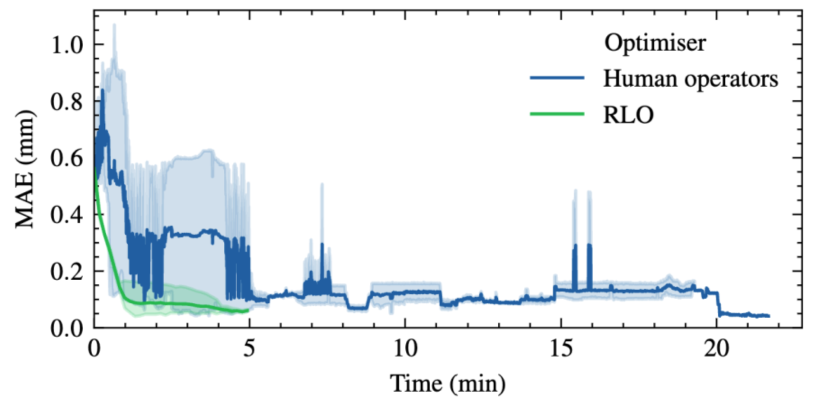
Reinforcement learning: From ARES Sinbad to the European XFEL

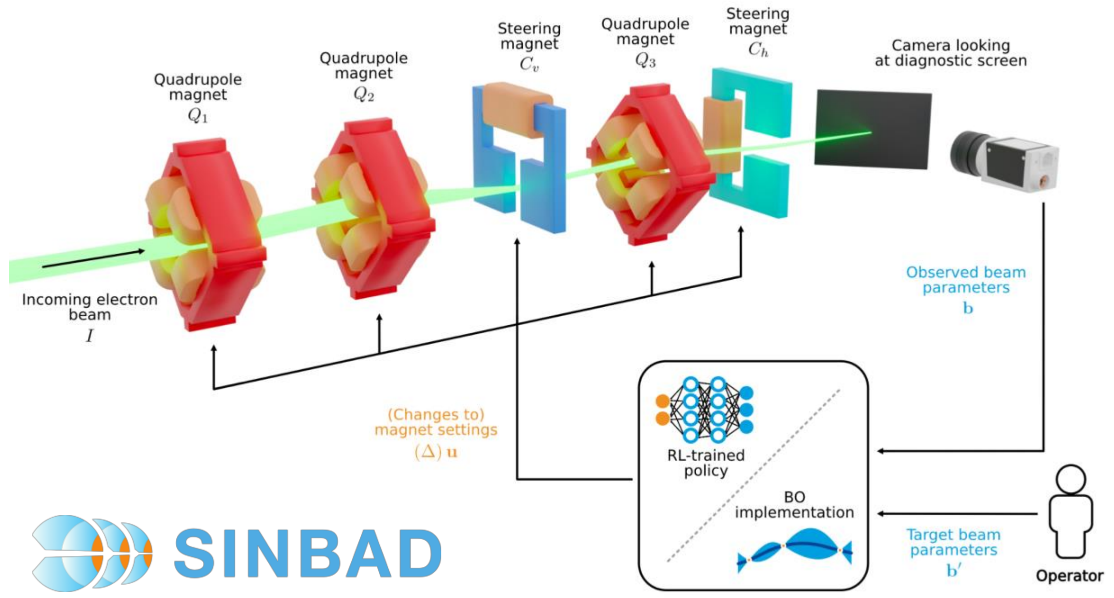
MATTER AND TECHNOLOGIES MT DESY.

Reinforcement learning-trained optimization at ARES

- Deploy a RL-trained optimization algorithm trained purely in simulation to the **real-world** with **zero-shot learning** thanks to **domain randomization.**
- The trained policy **outperforms other optimization algorithms and expert human operators**.







J Kaiser, O Stein, A Eichler. **Learning-based Optimisation of Particle Accelerators Under Partial Observability Without Real-World Training**. In *International Conference on Machine Learning*, 2022.
J Kaiser, C Xu, A Eichler, et. al. **Reinforcement learning-trained optimisers and Bayesian optimisation for online particle accelerator tuning**. In *Scientific reports* 14 (1), 2024

Cheetah: Speeding up simulations by 10^8

Linear beam dynamics simulation python package

MATTER AND TECHNOLOGIES MT DESY.

Why we need it? Training of RL agents would require **3 years of beam time** on the real machine, **11 days with Ocelot, 1 hour with Cheetah.**

Main features in support of ML applications:

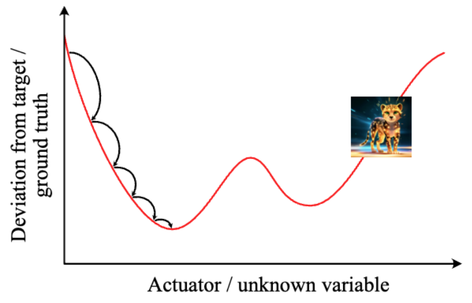
- **Ultra-fast** compute (at the cost of fidelity)
- **Differentiability**
- **GPU support**

TABLE I. Step computation times of simulation codes in milliseconds

Code	Comment	Laptop	HPC node
ASTRA	space charge	264 000.00	3 605 000.00
	no space charge	109 000.00	183 000.00
Parallel ASTRA	space charge	39 000.00	17 300.00
	no space charge	16 900.00	12 600.00
Ocelot	space charge	22 100.00	21 700.00
	no space charge	182.00	119.00
Bmad-X		40.50	74.30
Cheetah	ParticleBeam	0.79	0.72
	ParticleBeam + GPU	-	0.09
	ParameterBeam	0.02	0.04

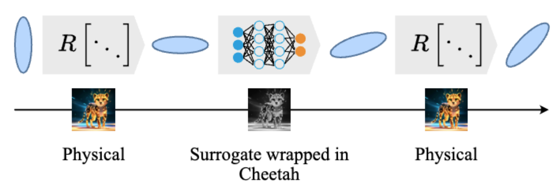
Making use of Cheetah's **differentiability**

- Bayesian optimization prior
- Gradient-based tuning / system identification



Making use of Cheetah's **speed**

- Reinforcement learning
- Integration of modular network surrogate



Cheetah in daily operation at LCLS:

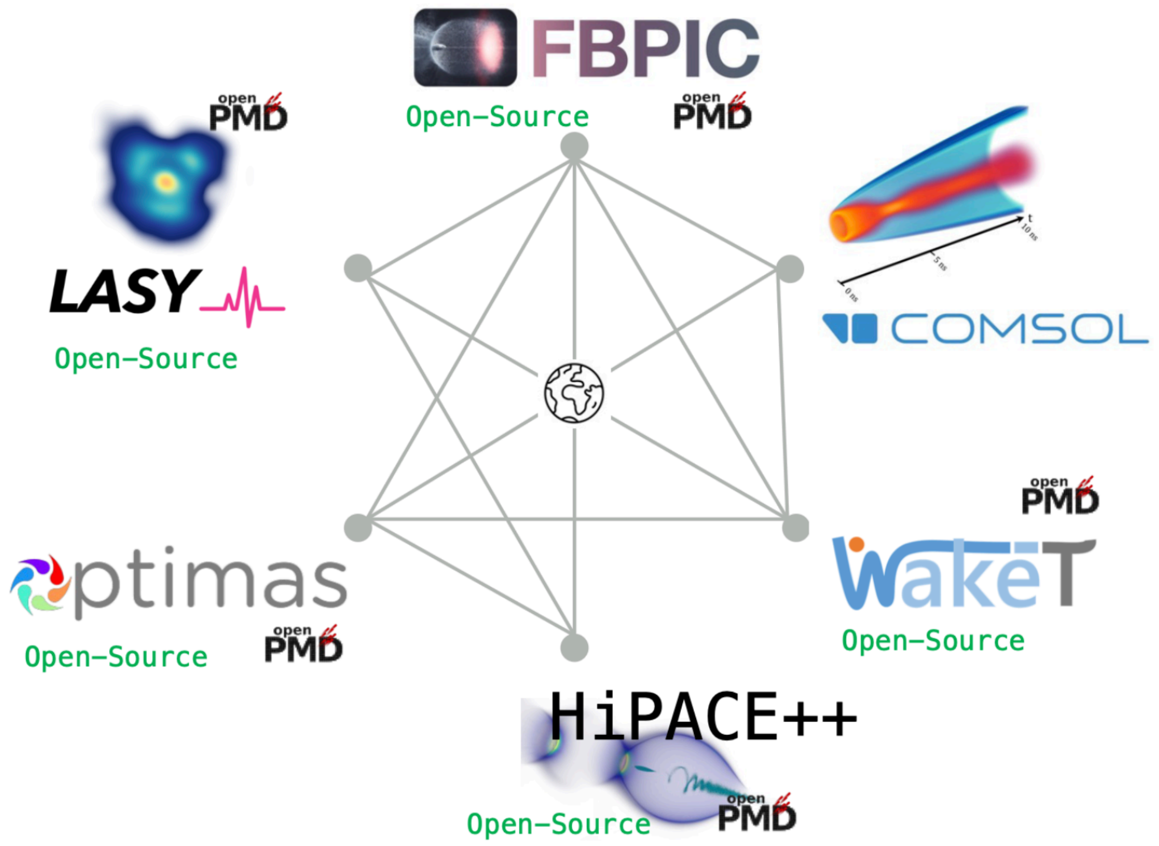
SLAC NATIONAL ACCELERATOR LABORATORY

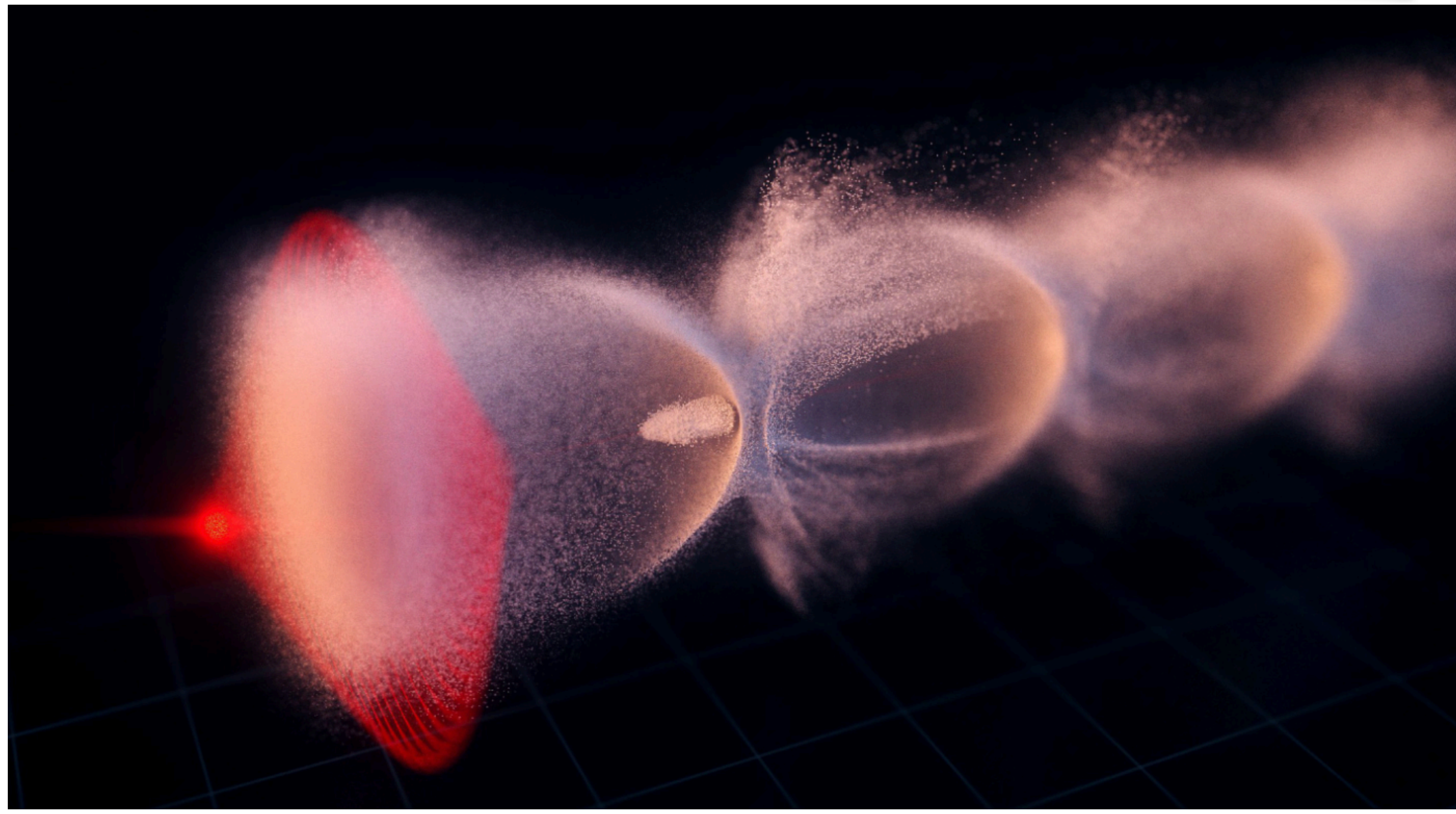
Now deployed to **daily** operations at **LCLS** for 6D-phase space reconstruction module.

Jan Kaiser, Chenran Xu, Annika Eichler and Andrea Santamaria Garcia. **Bridging the Gap Between Machine Learning and Particle Accelerator Physics with High-Speed, Differentiable Simulations.** In *Physical Review Accelerators and Beams*, 2024.

Similar pattern with plasma accelerators.

Opportunities with AI/ML surrogates of expensive Particle-in-Cell simulation?





Accelerating experimental design

MODE

JENAA
Joint ECFA-NuPECC-APPEC Activities

iris hep

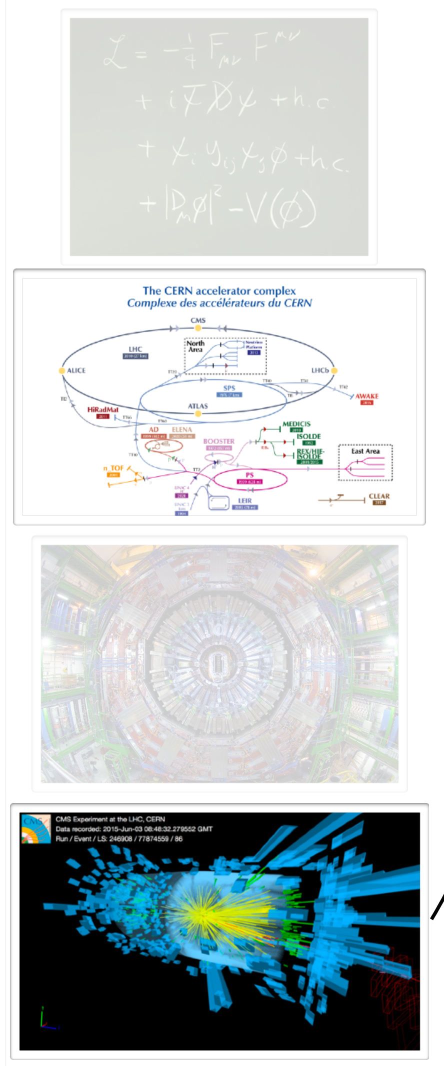
NSF

APPEC

Third MODE Workshop on
Differentiable Programming for Experiment Design

Princeton University
24-26 July, 2023

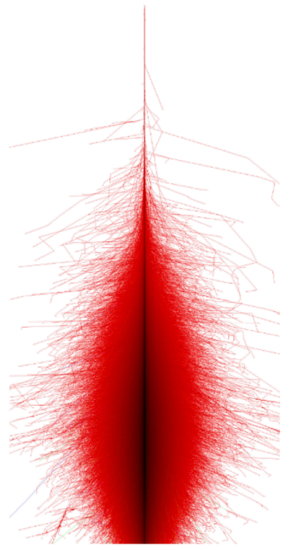
Experiment Design

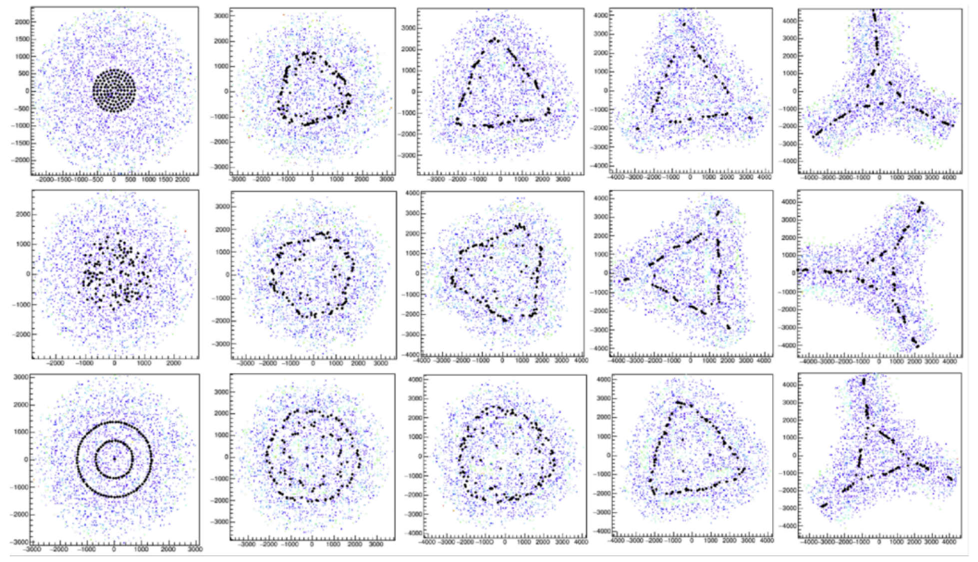


Experimental Design

Automatically learn to arrange sensors given a physics target

Example tuning positions of detectors for a gamma ray observatory





https://mode-collaboration.github.io/'; Dorigo et al 2310.01857

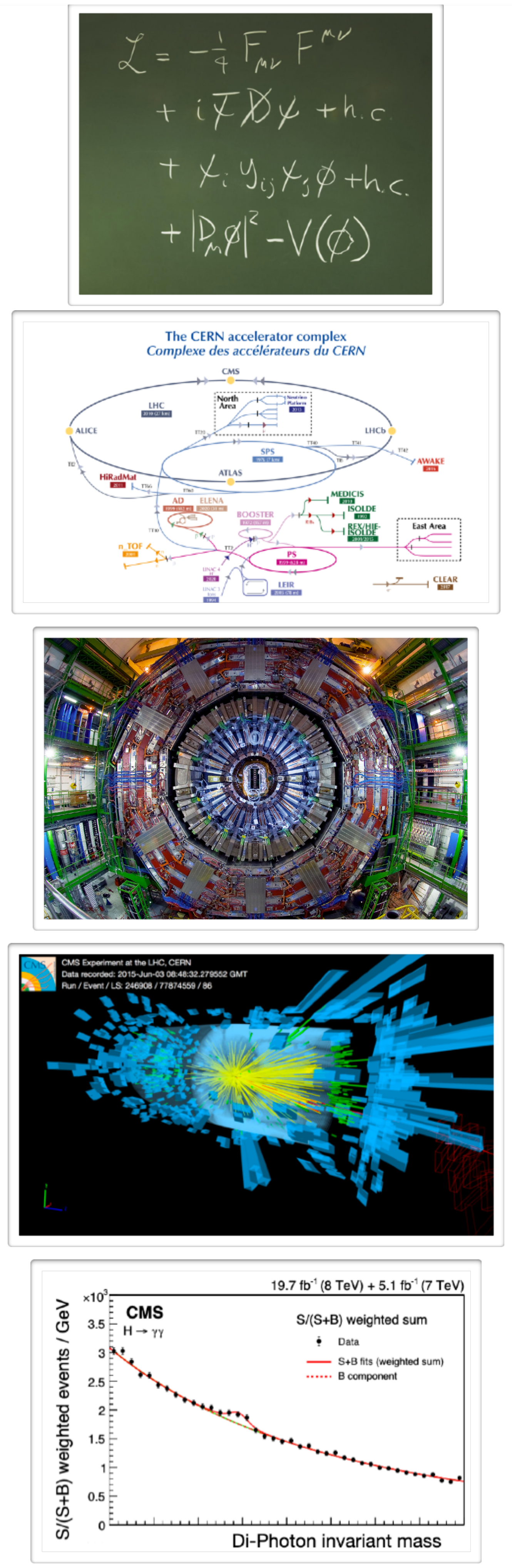
Differentiable versions of all steps in the particle physics processing chain

Either as ML-based surrogate models

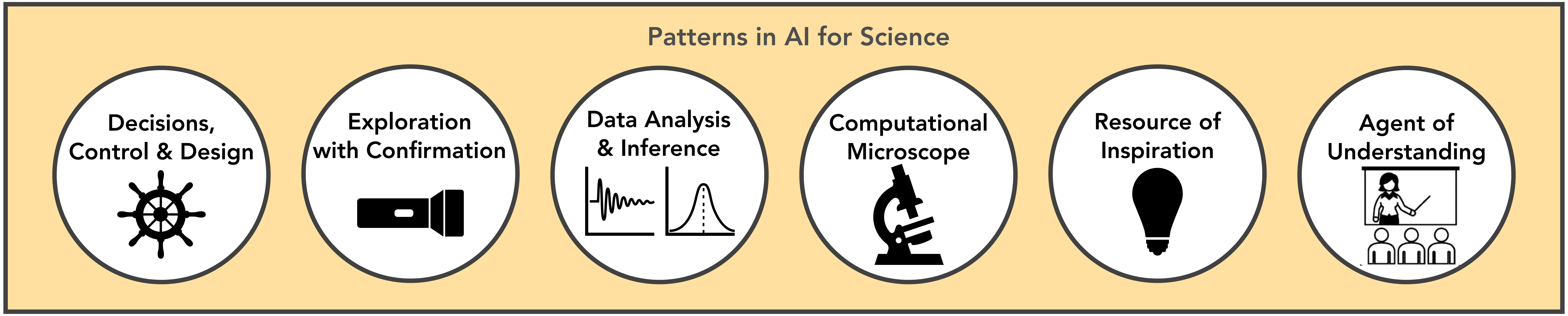
Or via e.g. differentiable programming

What can we do with this?

Heinrich, Kagan 2308.16680



Patterns in AI for Science



Adapted from figure by Mario Krenn, et. al. in https://arxiv.org/abs/2204.01467

Exploration with Confirmation

Drug & Materials Discovery

Many uses of AI aimed at accelerating drug discovery and materials discovery

- Experimental follow up needed to confirm the predicted properties
- OK if the predictions are wrong as long as it accelerates the discovery process

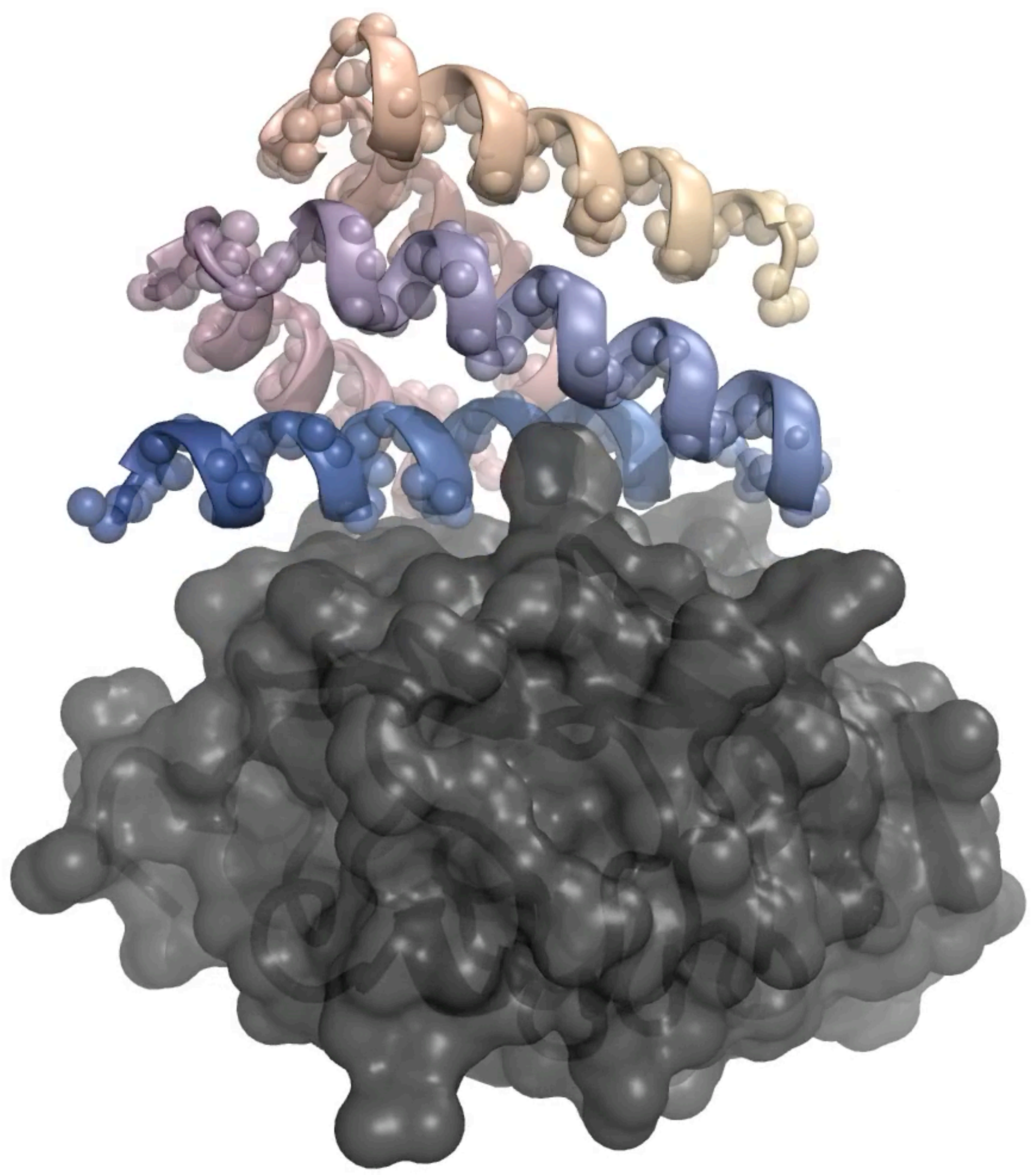
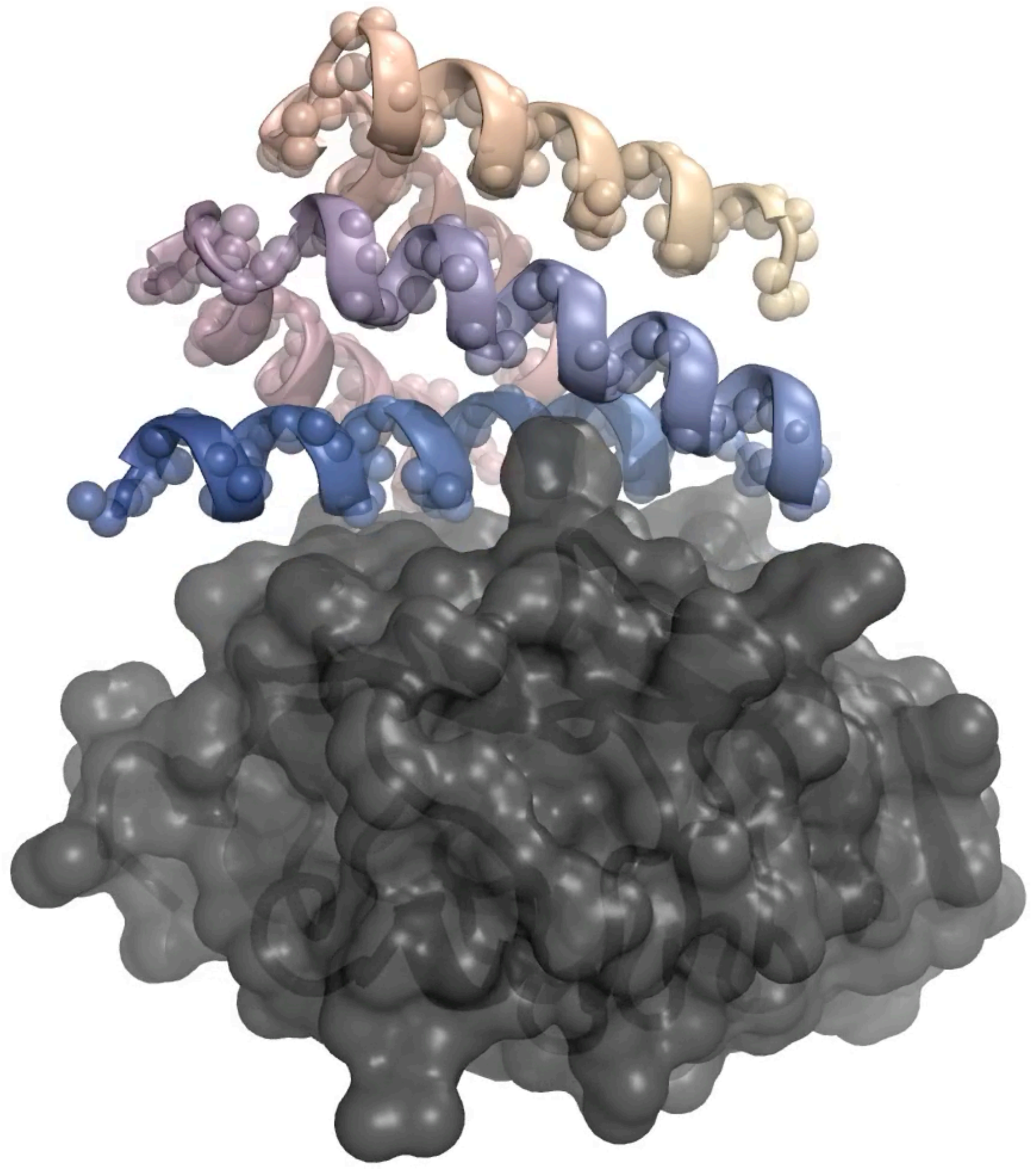


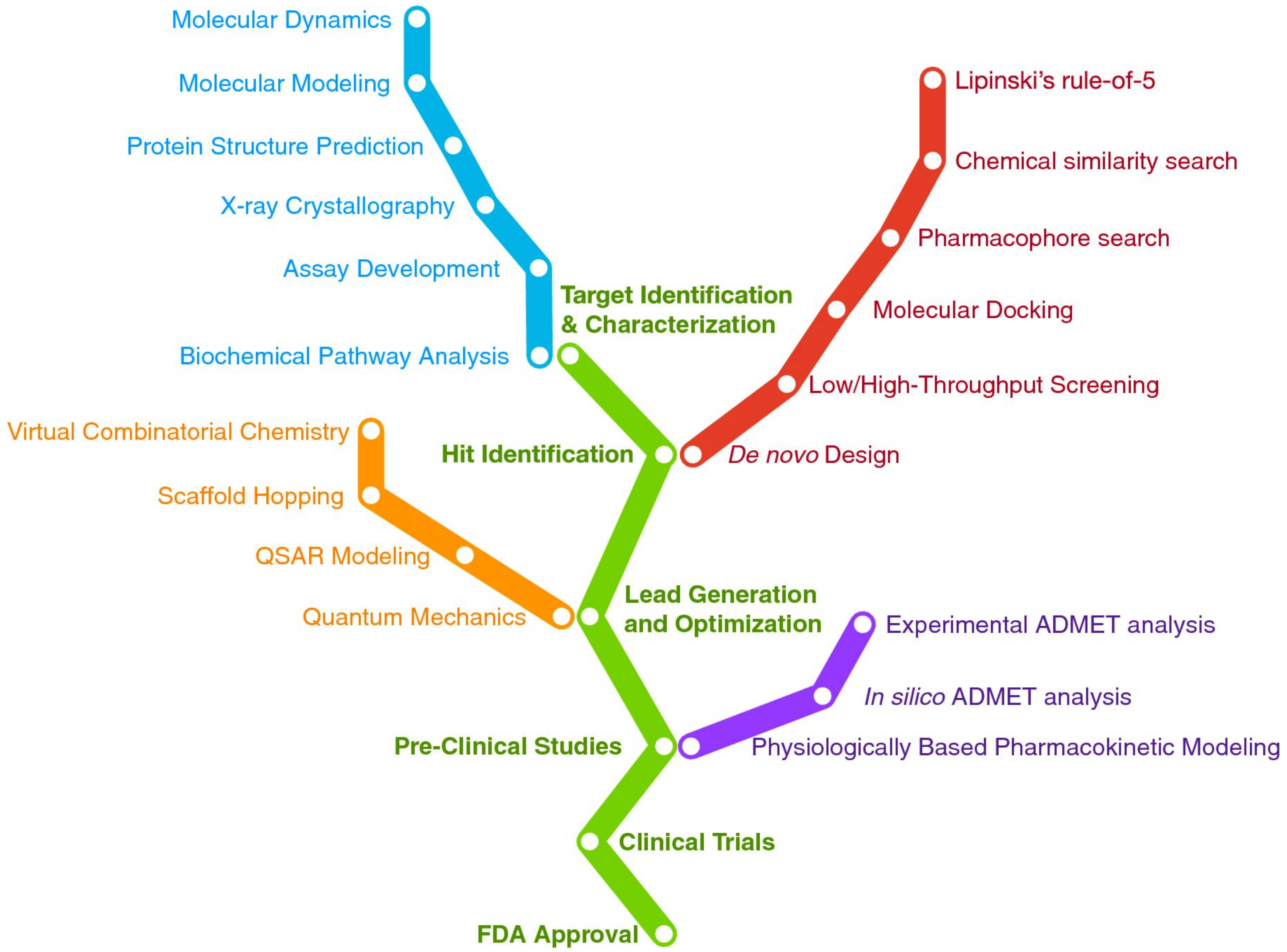
Image credits: RFDiffusion from Baker Lab, Institute for Protein Design, U Washington;
Chanin Nantasenamat from Towards reproducible computational drug discovery. *J Cheminform* **12**, 9 (2020). https://doi.org/10.1186/s13321-020-0408-x

https://dpmd.ai/PK-materials

Many uses of AI aimed at accelerating drug discovery and materials discovery

- Experimental follow up needed to confirm the predicted properties
- OK if the predictions are wrong as long as it accelerates the discovery process





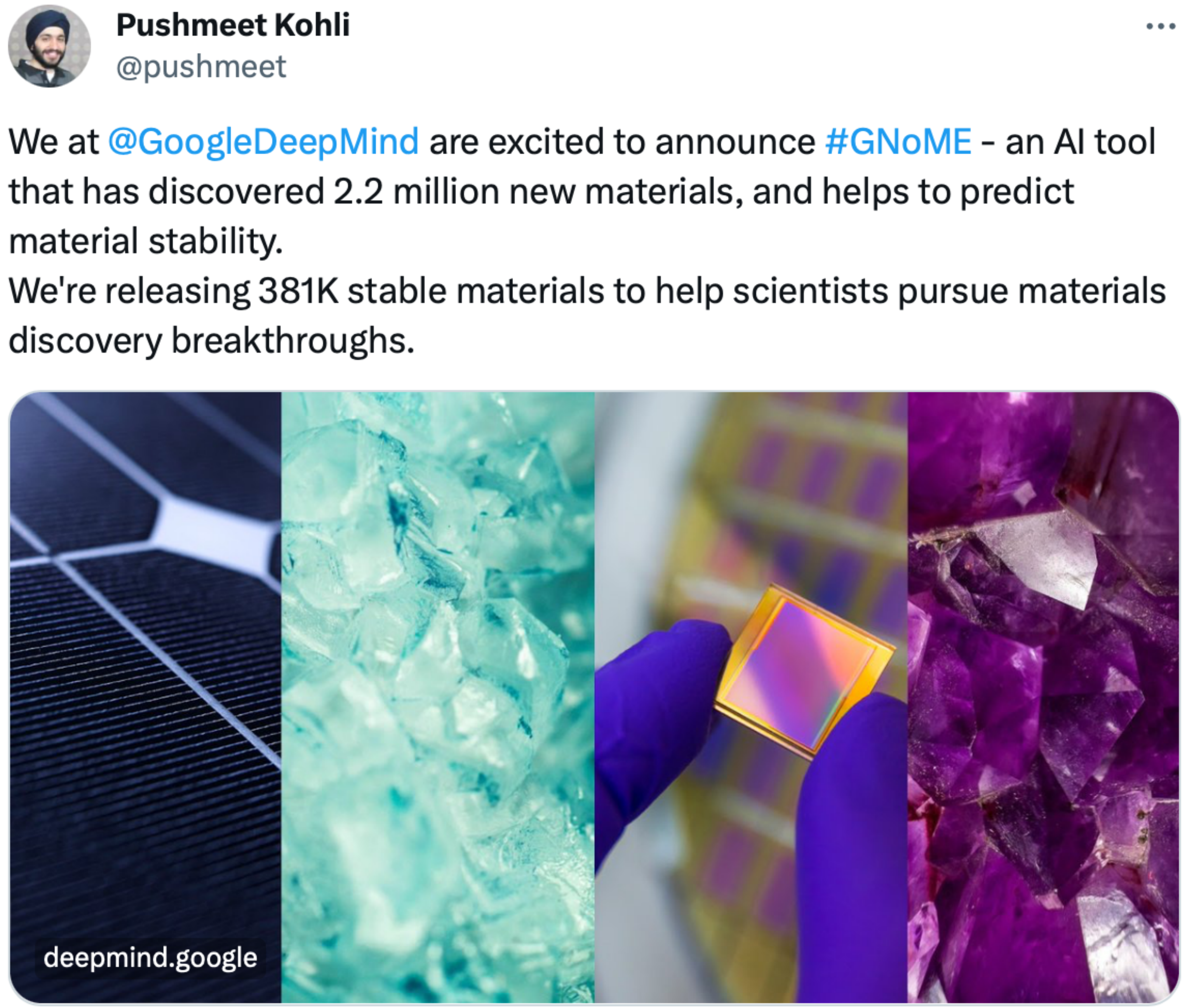


Image credits: RFDiffusion from Baker Lab, Institute for Protein Design, U Washington;
Chanin Nantasenamat from Towards reproducible computational drug discovery. *J Cheminform* **12**, 9 (2020). https://doi.org/10.1186/s13321-020-0408-x

https://dpmd.ai/PK-materials

Automated Theorem Proving

Theorem proving is another example of exploration paired with confirmation

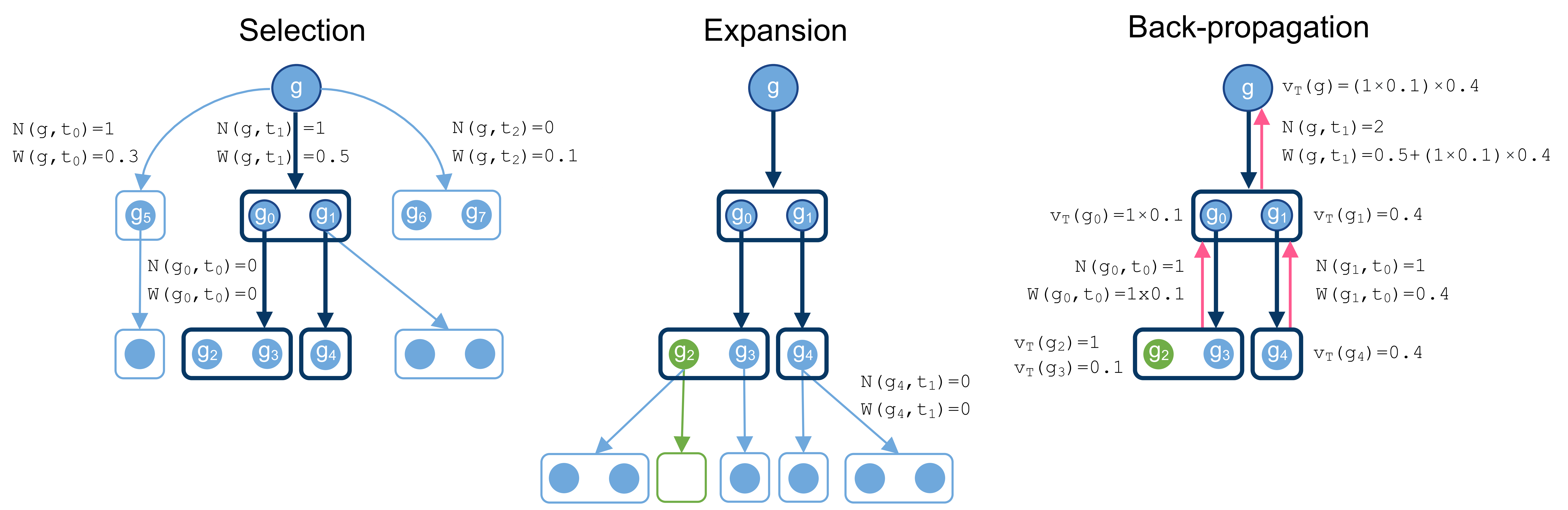
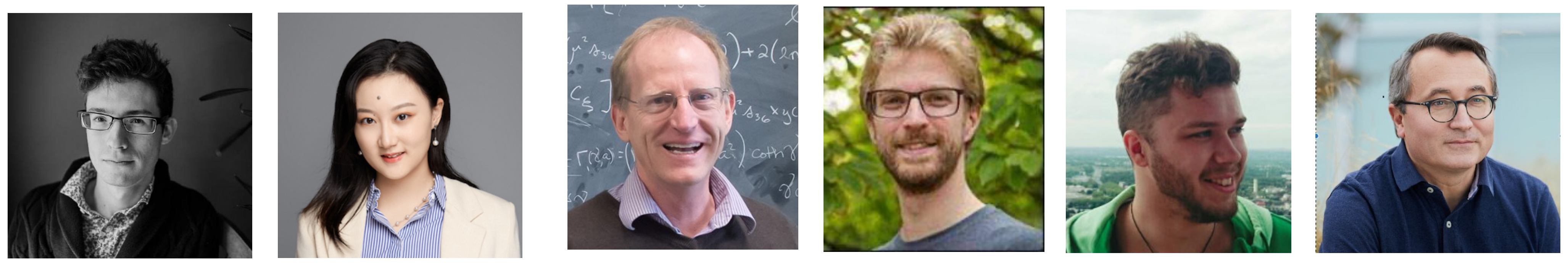


Figure 5: **HyperTree Proof Search**. We aim at finding a proof of the root theorem g with HTPS. Proving either $\{g_5\}$, $\{g_0, g_1\}$, or $\{g_6, g_7\}$ would lead to a proof of g by tactic t_0, t_1, or t_2. The figure represents the three steps of HTPS that are repeated until a proof is found. Guided by the search policy, we select a hypertree whose leaves are unexpanded nodes. The selected nodes are then expanded, adding new tactics and nodes to the hypergraph. Finally, during back-propagation we evaluate the node values of the hypertree, starting from the leaves back to the root, and update the visit counts and total action values.

AI for Amplitudes



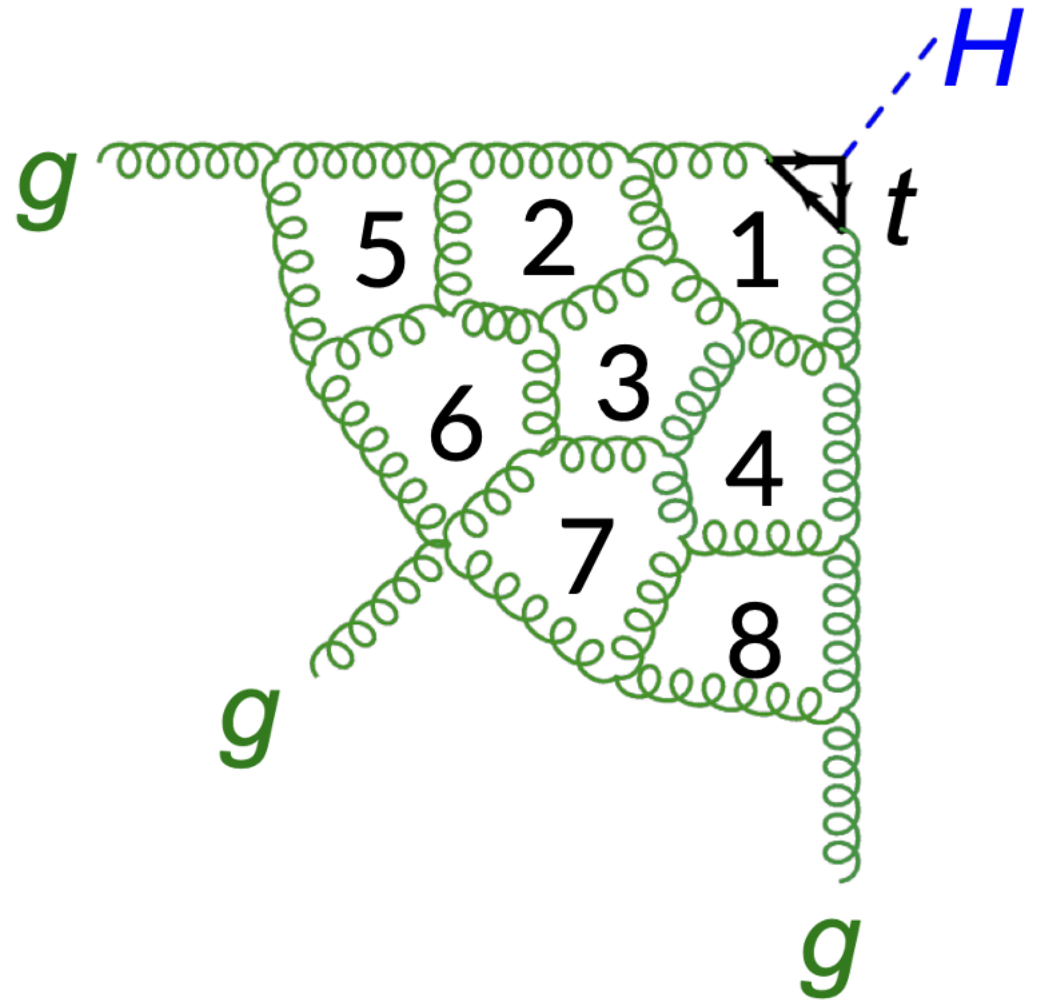
Similarly, we are using generative AI to help compute multi-loop scattering amplitudes

- **the answer is hard to find, but easy to check.**

We don't need the model to be provably correct, we just need it to be good at guessing because we can get a certificate of correctness

- The problem is inherently discrete, so transformers are a natural choice

We see ~99% accuracy in predicting the coefficients of the amplitude!



MACHINE LEARNING
Science and Technology

PAPER

Transforming the bootstrap: using transformers to compute scattering amplitudes in planar $\mathcal{N} = 4$ super Yang–Mills theory

Tianji Cai[1,5,*], Garrett W Merz[2,5,*], François Charton[3,5], Niklas Nolte[3], Matthias Wilhelm[4], Kyle Cranmer[2] and Lance J Dixon[1]

Meta §

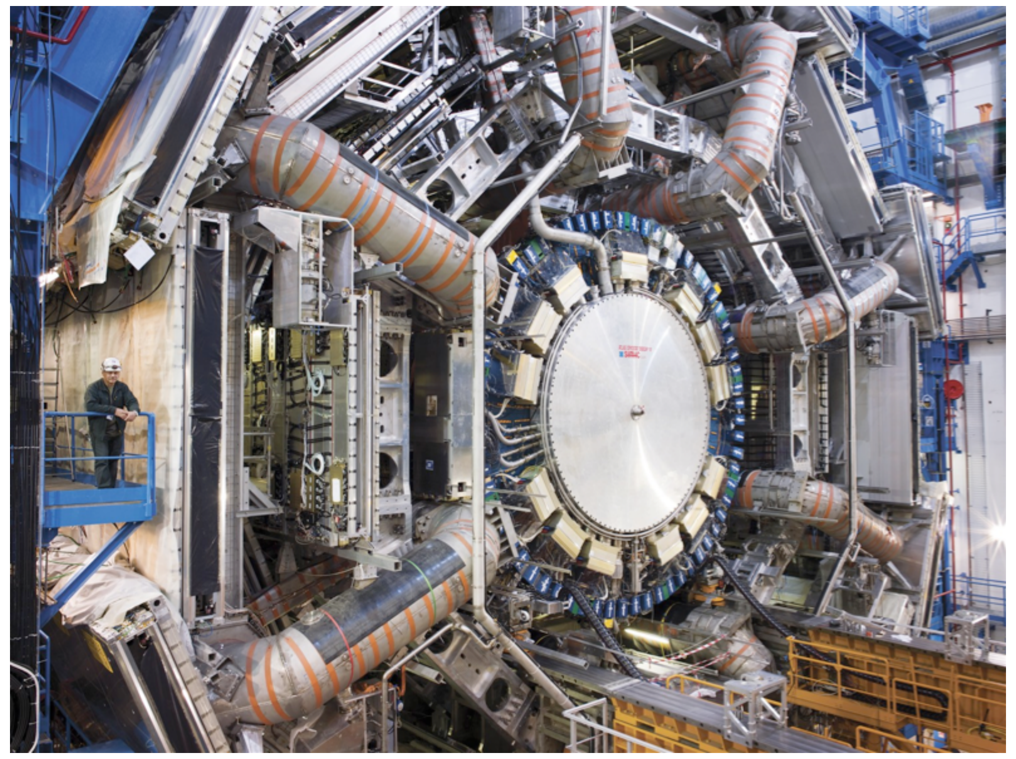
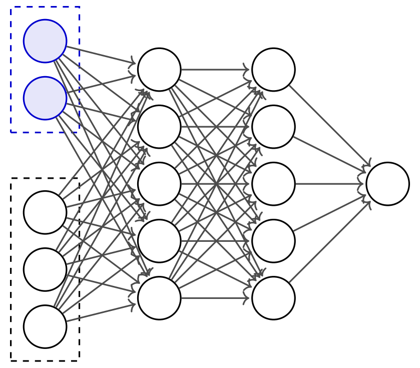
WISCONSIN † UNIVERSITY OF WISCONSIN–MADISON

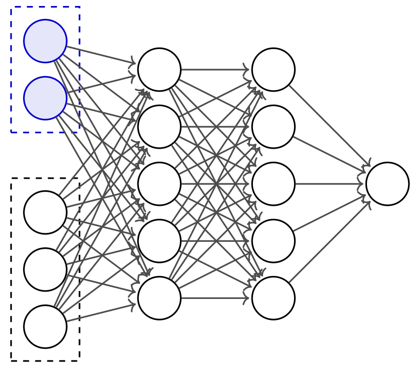
UNIVERSITY OF COPENHAGEN ‡

SLAC NATIONAL ACCELERATOR LABORATORY

In contrast, AI/ML in experimental physics, astrophysics, and cosmology is often a component of a **hypothesis testing / statistical inference pipeline**.

- Robustness to systematic uncertainty (distribution shift) is important!
- Mistakes matter — we need to be able to calibrate & perform uncertainty quantification!




Data Collection | Data Analysis Pipeline | Scientific Claims

Simulation-based Inference

Statistical Framing

forward modeling
generation
simulation

PREDICTION

$p(x, z \mid \theta)$

θ
parameters of interest

z
latent variables

x
observed data
simulated data

INFERENCE

inverse problem
measurement
parameter estimation

Simulating particle physics processes

Theory
parameters
θ

Evolution

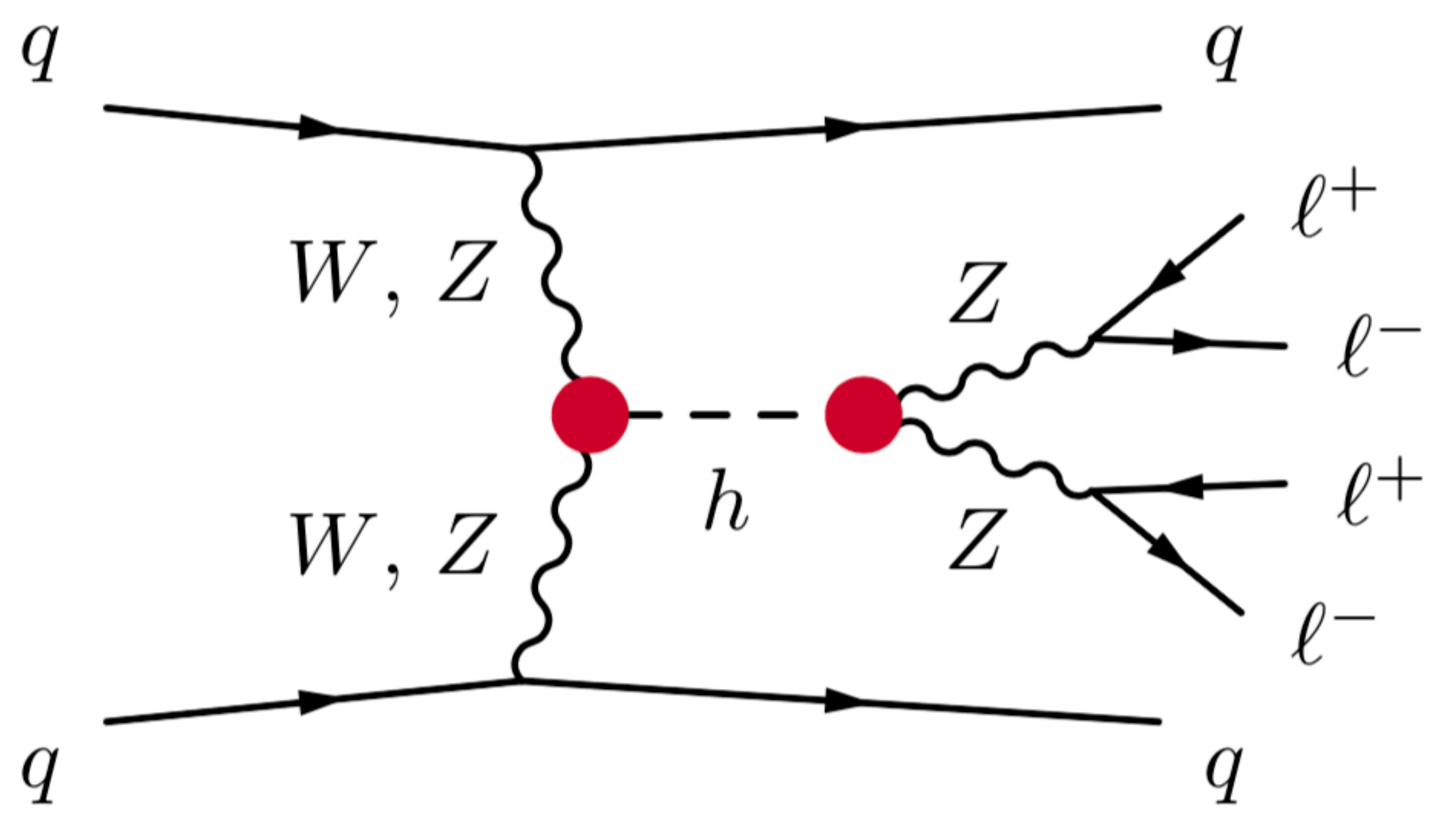
Simulating particle physics processes

Latent variables

Parton-level momenta

Theory parameters

$$z_p \longleftarrow \theta$$

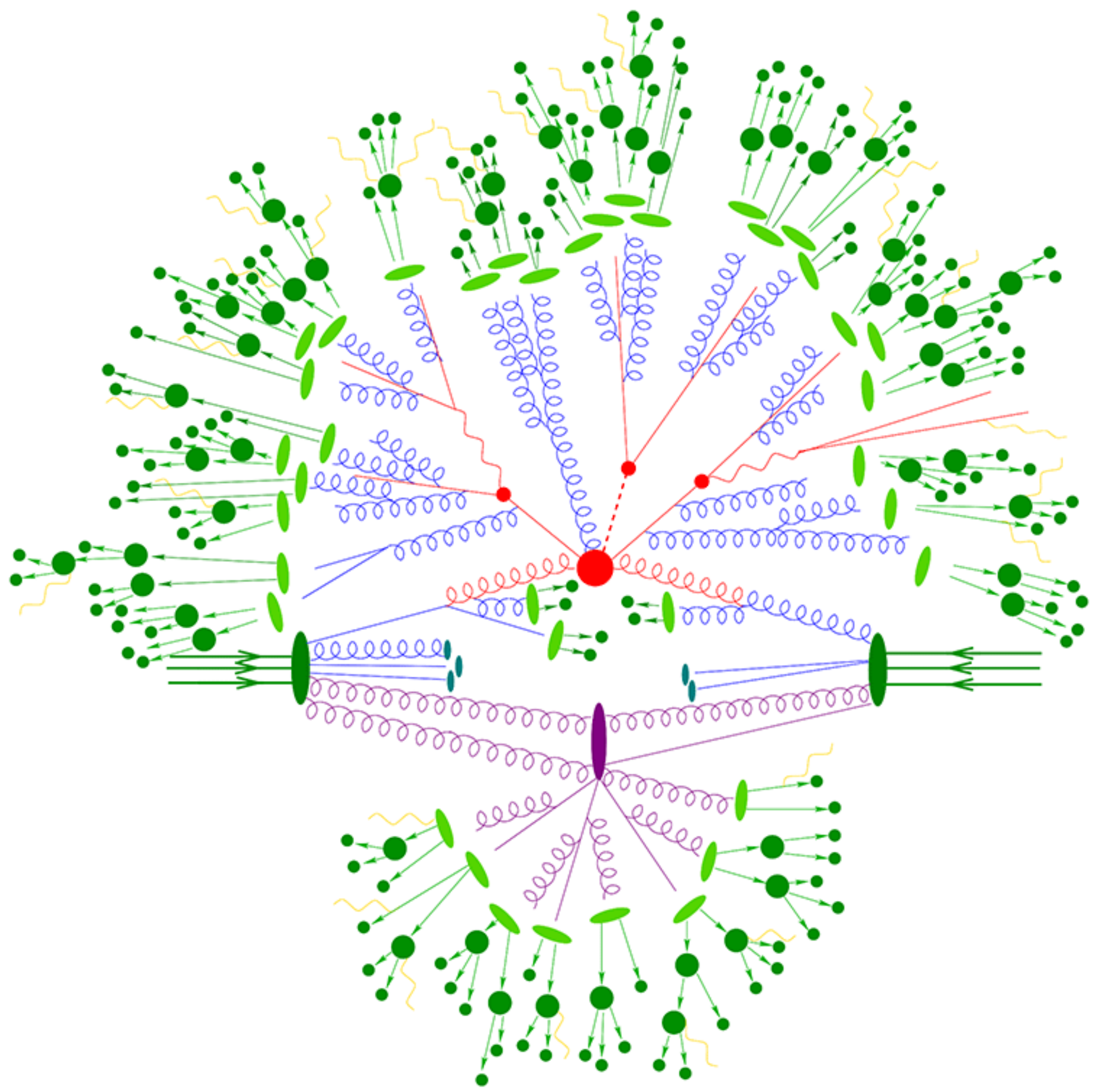


Evolution

Simulating particle physics processes

Latent variables

Shower splittings		Parton-level momenta		Theory parameters
z_s	$\leftarrow$	z_p	$\leftarrow$	θ



[F. Krauss]

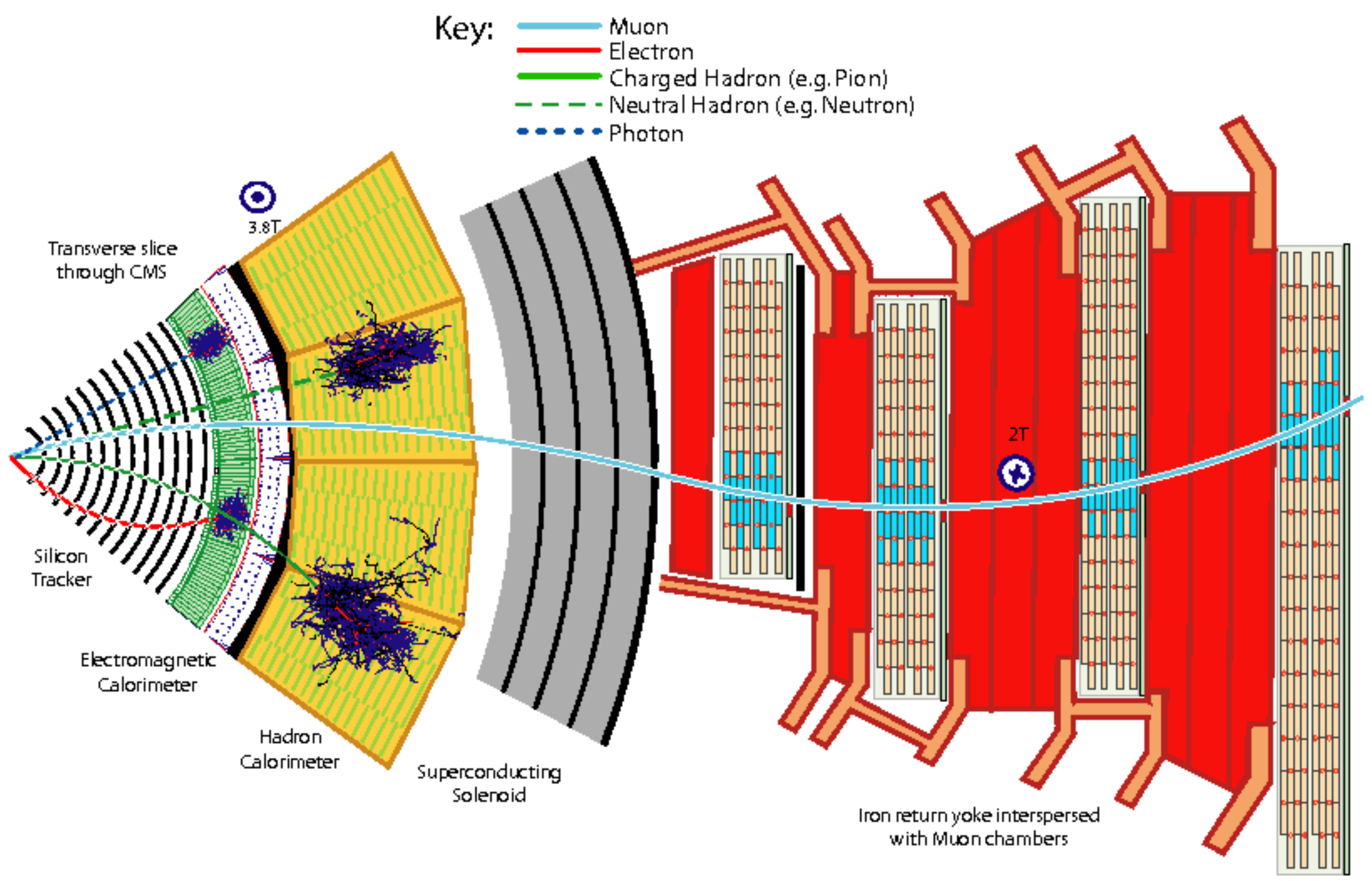
$\longleftarrow$ Evolution

Simulating particle physics processes

Latent variables

Detector interactions	Shower splittings	Parton-level momenta	Theory parameters

$$z_d \longleftarrow z_s \longleftarrow z_p \longleftarrow \theta$$



[CMS]

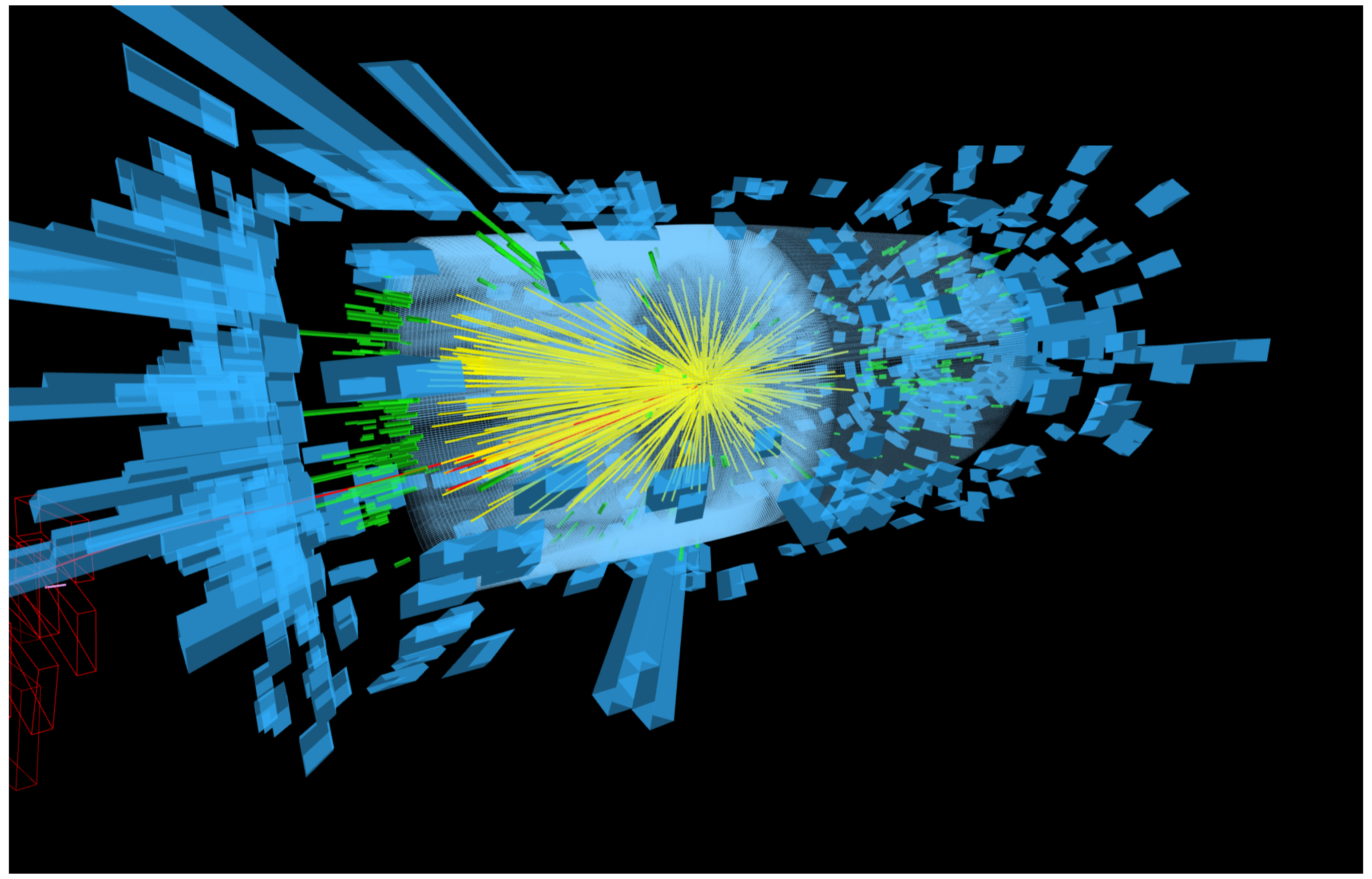
Evolution

Simulating particle physics processes

Latent variables

Observables		Detector interactions	Shower splittings	Parton-level momenta	Theory parameters

$$x \leftarrow z_d \leftarrow z_s \leftarrow z_p \leftarrow \theta$$



Evolution

Simulating particle physics processes

Latent variables

Observables	Detector interactions	Shower splittings	Parton-level momenta	Theory parameters
x	z_d	z_s	z_p	θ

$$x \longleftarrow z_d \longleftarrow z_s \longleftarrow z_p \longleftarrow \theta$$

Sample from $p(x|z_d)$ $p(z_d|z_s)$ $p(z_s|z_p)$ $p(z_p|\theta)$

GEANT4
A SIMULATION TOOLKIT

MADGRAPH5_aMC@NLO

Prediction (simulation)

Simulating particle physics processes

Latent variables

Observables	Detector interactions	Shower splittings	Parton-level momenta	Theory parameters

$$x \longleftarrow z_d \longleftarrow z_s \longleftarrow z_p \longleftarrow \theta$$

$$p(x|\theta) = \int \mathrm{d}z_d \int \mathrm{d}z_s \int \mathrm{d}z_p \quad p(x|z_d) \qquad p(z_d|z_s) \qquad p(z_s|z_p) \qquad p(z_p|\theta)$$

$\longrightarrow$

Inference

Simulating particle physics processes

Latent variables

Observables | Detector interactions | Shower splittings | Parton-level momenta | Theory parameters

$$x \longleftarrow z_d \longleftarrow z_s \longleftarrow z_p \longleftarrow \theta$$

$$p(x|\theta) = \boxed{\int \mathrm{d}z_d \int \mathrm{d}z_s \int \mathrm{d}z_p} \; p(x|z_d) \quad p(z_d|z_s) \quad p(z_s|z_p) \quad p(z_p|\theta)$$

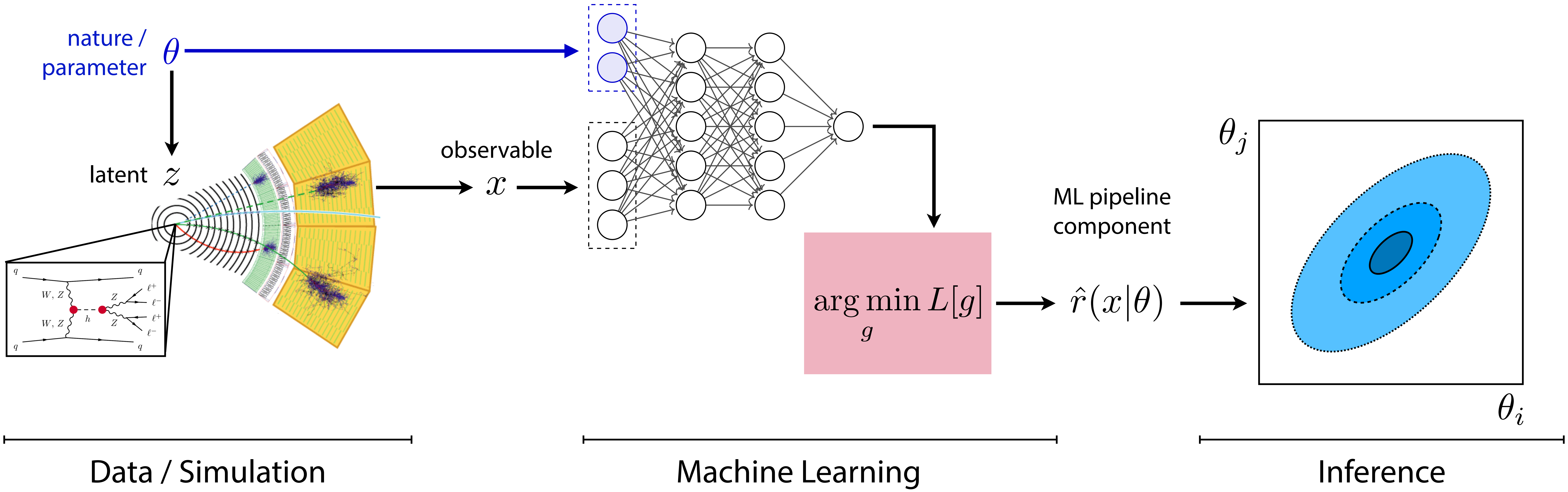
It's infeasible to calculate the integral over this enormous space!

Inference

Simulation-Based Inference

Deep learning and neural density estimation are effective at learning approximate surrogates for the fully differential likelihood (and posterior).
This is **revolutionizing principled statistical inference in science!**

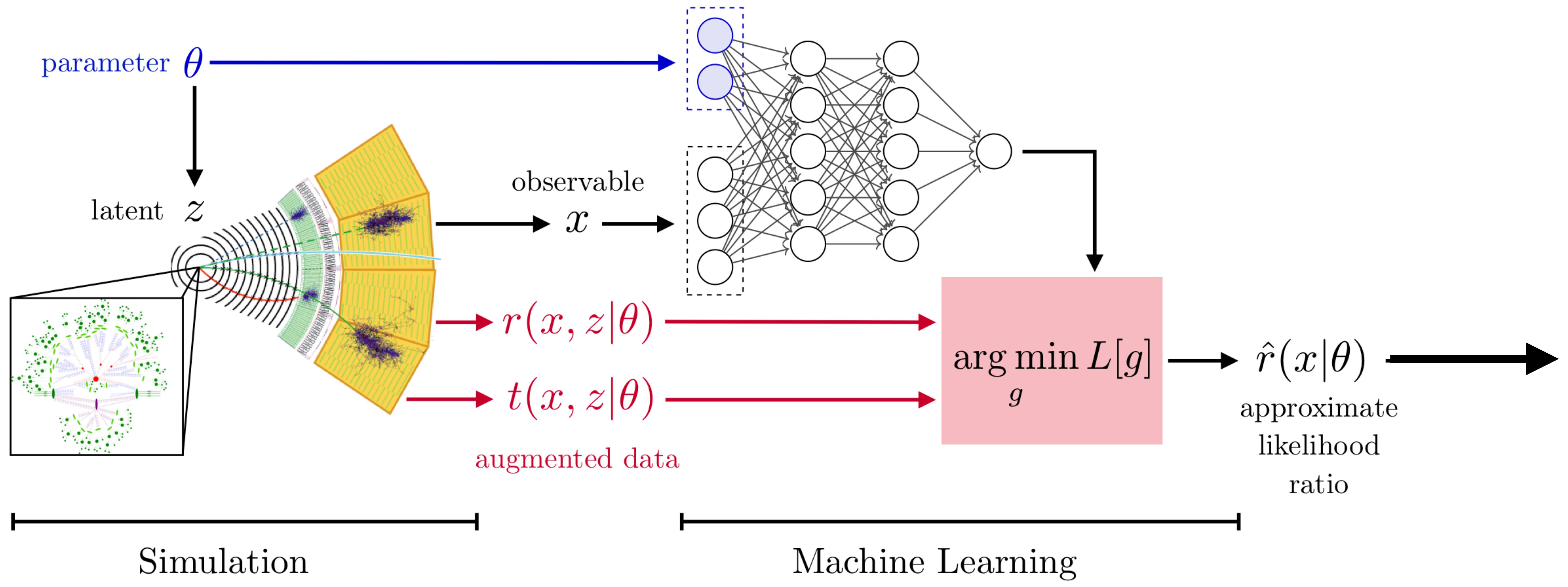
- Removes the need for hand-engineered summary statistics that sacrifice power

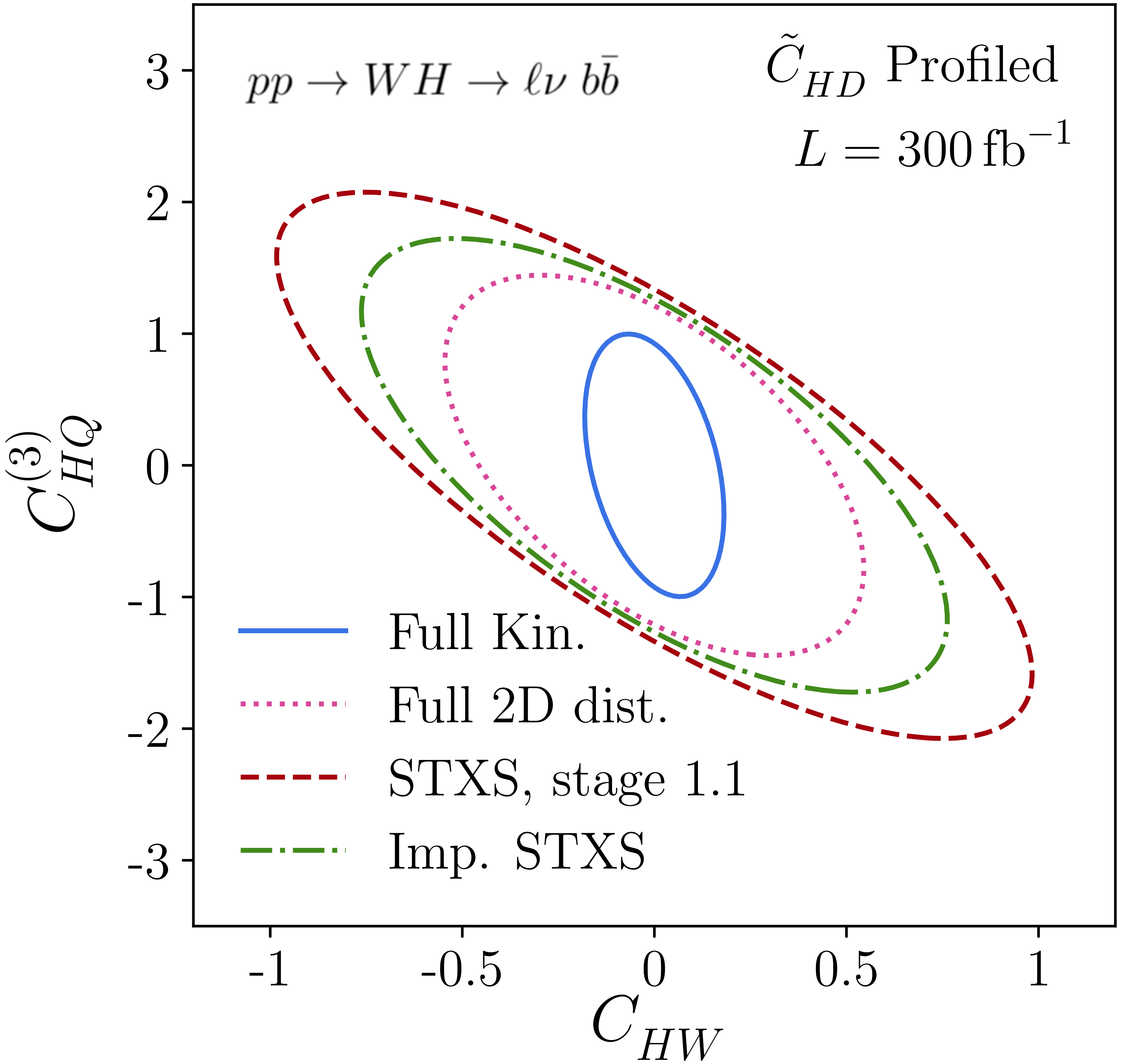


Impact on Studies of The Higgs Boson

Potential for massive gains in precision of a flagship measurement at the LHC !

Equivalent increasing data collected by LHC by several factors





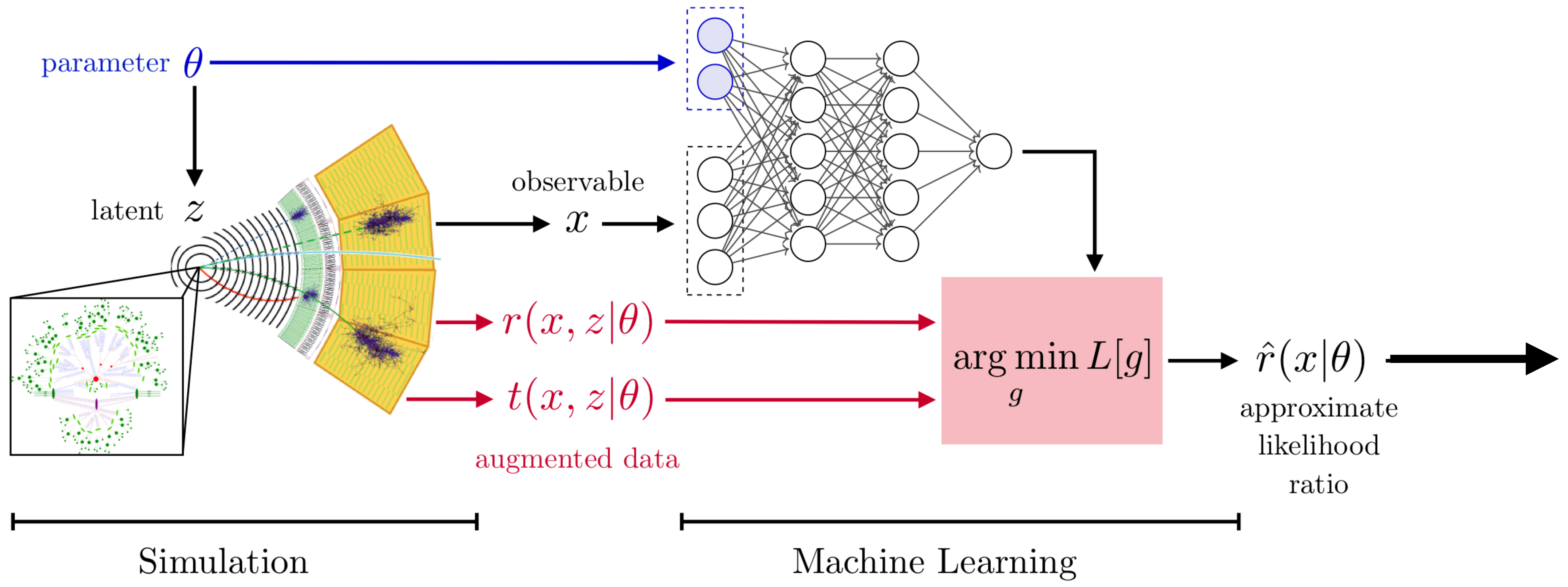
[J. Brehmer, S. Dawson, S. Homiller, F. Kling, T. Plehn 1908.06980]

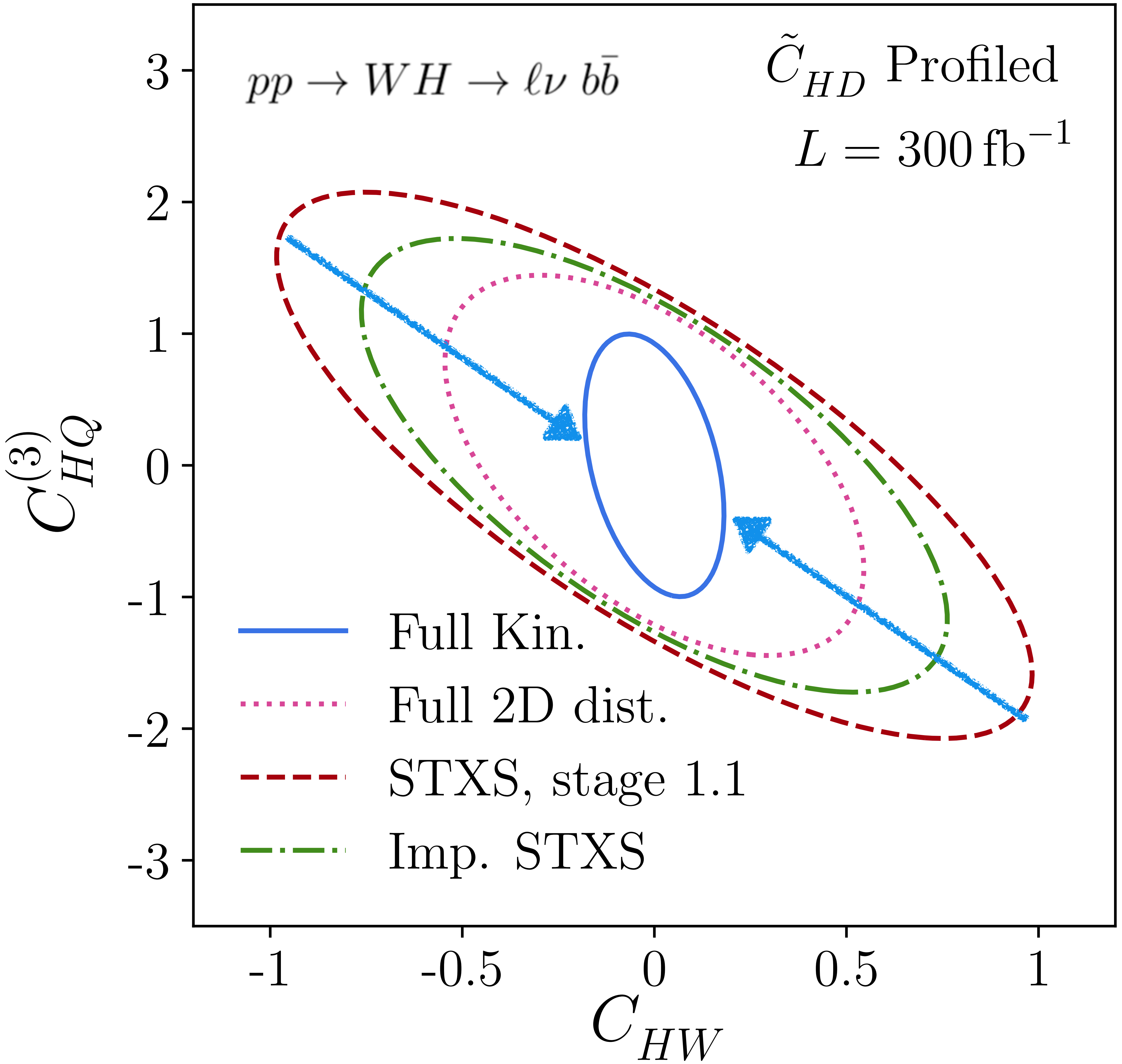
[J. Brehmer, F. Kling, I. Espejo, K. Cranmer 1907.10621]

Impact on Studies of The Higgs Boson

Potential for massive gains in precision of a flagship measurement at the LHC !

Equivalent increasing data collected by LHC by several factors





[J. Brehmer, S. Dawson, S. Homiller, F. Kling, T. Plehn 1908.06980]

[J. Brehmer, F. Kling, I. Espejo, K. Cranmer 1907.10621]


MAX-PLANCK-INSTITUT FÜR PHYSIK

SBI at EuCAIFCon 2024

Simulation-based Inference was well represented at EuCAIFCon2024!

Simulation of Z2 model using Variational Autoregressive Network (VAN). *Vaibhav Chahar* *UvA 1, Hotel CASA* 13:30 - 13:33	**Characterizing the Fermi-LAT high-latitude sky with simulation-based inference** *Christopher Eckner* *Sorbonne, Hotel CASA* 14:50 - 14:53	**Analyzing ML-enabled Full Population Model for Galaxy SEDs with Unsupervised Learning and Mutual Information** *Dr Sinan Deger*
Artificial Intelligence techniques in KM3NeT *Evangelia Drakopoulou* *UvA 1, Hotel CASA* 13:33 - 13:36	**Simulation-Based Supernova Ia Cosmology** *Konstantin Karchev* *Sorbonne, Hotel CASA* 14:53 - 14:56	**Convolutional neural network search for long-duration transient gravitational waves from glitching pulsars** *Rodrigo Tenorio*
ML-based Unfolding Techniques for High Energy Physics *Nathan Huetsch* *UvA 1, Hotel CASA* 13:36 - 13:39	**Optimizing bayesian inference in cosmology with Marginal Neural Ratio Estimation** *Guillermo Franco Abellan* *Sorbonne, Hotel CASA* 14:56 - 14:59	**Tuning neural posterior estimation for gravitational wave inference** *Alex Kolmus* *Oxford, Hotel CASA* 16:06 - 16:09
Building sparse kernel methods via dictionary learning. Expressive, regularized and interpretable models for statistical *Gaia Grosso*	**Stochastic Gravitational Wave Background Analysis with SBI** *James Alvey* *Sorbonne, Hotel CASA* 14:59 - 15:02	**Normalising flows for dense matter equation of state inference from gravitational wave observations of neutron star me** *Jessica Irwin*
pop-cosmos: comprehensive forward modelling of photometric galaxy survey data *Stephen Thorp* *UvA 1, Hotel CASA* 13:42 - 14:02	**COSMOPOWER: fully-differentiable Bayesian cosmology with neural emulators** *Alessio Spurio Mancini* *Sorbonne, Hotel CASA* 15:02 - 15:22	**A Strong Gravitational Lens Is Worth a Thousand Dark Matter Halos: Inference on Small-Scale Structure Using Sequent** *Sebastian Wagner-Carena*
Calibrating Bayesian Tension Statistics with Neural Ratio Estimators *Harry Bevins* *UvA 1, Hotel CASA* 14:02 - 14:22	**Networks Learning the Universe: From 3D (cosmological inference) to 1D (classification of spectra)** *Caroline Heneka* *Sorbonne, Hotel CASA* 15:22 - 15:42	**Optimal, fast, and robust inference of reionization-era cosmology with the 21cmPIE-INN** *Benedikt Schosser* *Oxford, Hotel CASA* 16:49 - 16:52
Machine learning for radiometer calibration in global 21cm cosmology *Mr Samuel Alan Kossoff Leeney* *UvA 1, Hotel CASA* 14:22 - 14:25	**Anomaly aware machine learning for dark matter direct detection at DARWIN** *Andre Scaffidi* *Sorbonne, Hotel CASA* 15:42 - 15:45	**Simulation Based Inference from the CD-EoR 21-cm signal** *Anchal Saxena* *Oxford, Hotel CASA* 16:52 - 16:55
PolySwyft: a sequential simulation-based nested sampler *Kilian Scheutwinkel* *UvA 1, Hotel CASA* 14:28 - 14:31	**Clustering Considerations for Nested Sampling** *Adam Ormondroyd* *Sorbonne, Hotel CASA* 15:45 - 15:48	**Flexible conditional normalizing flow distributions over manifolds: the jammy-flows toolkit** *Dr Thorsten Glüsenkamp* *Oxford, Hotel CASA* 16:55 - 16:58
Extracting Dark Matter Halo Parameters with Overheated Exoplanets *María Benito* *UvA 1, Hotel CASA* 14:31 - 14:34	**Enhancing Robustness: BSM Parameter Inference with n1D-CNN and Novel Data Augmentation** *Yong Sheng Koay* *Sorbonne, Hotel CASA* 15:48 - 15:51	**A deep learning method for the trajectory reconstruction of gamma rays with the DAMPE space mission** *Parzival Nussbaum*
	Fully Bayesian Forecasts with Neural Bayes Ratio Estimation *Thomas Gessey-Jones* *Sorbonne, Hotel CASA* 15:51 - 15:54	

Summary talks: Astroparticle Physics and AI (Siddarth Mishra-Sharma) *Tilman Plehn*
UvA 2-3-4, Hotel CASA 09:00 - 09:40

Summary talks: Cosmology and AI (Benjamin Wandelt) *David Rousseau*
UvA 2-3-4, Hotel CASA 15:00 - 15:40

2024: https://indico.nikhef.nl/event/4875/
2025: https://agenda.infn.it/event/43565/

First LHC papers using Simulation-Based Inference

EUROPEAN ORGANISATION FOR NUCLEAR RESEARCH (CERN)

Submitted to: Rep. Prog. Phys.

CERN-EP-2024-298
December 3, 2024

Measurement of off-shell Higgs boson production in the $H^* \to ZZ \to 4\ell$ decay channel using a neural simulation-based inference technique in 13 TeV pp collisions with the ATLAS detector

The ATLAS Collaboration

A measurement of off-shell Higgs boson production in the $H^* \to ZZ \to 4\ell$ decay channel is presented. The measurement uses 140 fb^{-1} of proton–proton collisions at $\sqrt{s} = 13$ TeV collected by the ATLAS detector at the Large Hadron Collider and supersedes the previous result in this decay channel using the same dataset. The data analysis is performed using a neural simulation-based inference method, which builds per-event likelihood ratios using neural networks. The observed (expected) off-shell Higgs boson production signal strength in the $ZZ \to 4\ell$ decay channel at 68% CL is $0.87^{+0.75}_{-0.54}$ ($1.00^{+1.04}_{-0.95}$). The evidence for off-shell Higgs boson production using the $ZZ \to 4\ell$ decay channel has an observed (expected) significance of 2.5σ (1.3σ). The expected result represents a significant improvement relative to that of the previous analysis of the same dataset, which obtained an expected significance of 0.5σ. When combined with the most recent ATLAS measurement in the $ZZ \to 2\ell 2\nu$ decay channel, the evidence for off-shell Higgs boson production has an observed (expected) significance of 3.7σ (2.4σ). The off-shell measurements are combined with the measurement of on-shell Higgs boson production to obtain constraints on the Higgs boson total width. The observed (expected) value of the Higgs boson width at 68% CL is $4.3^{+2.7}_{-1.9}$ ($4.1^{+3.5}_{-3.4}$) MeV.

EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH (CERN)

CERN-EP-2024-294
2024/11/27

CMS-HIG-23-016

Constraints on standard model effective field theory for a Higgs boson produced in association with W or Z bosons in the $\mathrm{H} \to \mathrm{b\bar{b}}$ decay channel in proton-proton collisions at $\sqrt{s} = 13\,\mathrm{TeV}$

The CMS Collaboration*

Abstract

A standard model effective field theory (SMEFT) analysis with dimension-six operators probing nonresonant new physics effects is performed in the Higgs-strahlung process, where the Higgs boson is produced in association with a W or Z boson, in proton-proton collisions at a center-of-mass energy of 13 TeV. The final states in which the W or Z boson decays leptonically and the Higgs boson decays to a pair of bottom quarks are considered. The analyzed data were collected by the CMS experiment between 2016 and 2018 and correspond to an integrated luminosity of 138 fb^{-1}. An approach designed to simultaneously optimize the sensitivity to Wilson coefficients of multiple SMEFT operators is employed. Likelihood scans as functions of the Wilson coefficients that carry SMEFT sensitivity in this final state are performed for different expansions in SMEFT. The results are consistent with the predictions of the standard model.

Submitted to the Journal of High Energy Physics

First LHC papers using Simulation-Based Inference

EUROPEAN ORGANISATION FOR NUCLEAR RESEARCH (CERN)

Submitted to: Rep. Prog. Phys.

CERN-EP-2024-298
December 3, 2024

Measurement of off-shell Higgs boson production in the $H^* \to ZZ \to 4\ell$ decay channel using a neural simulation-based inference technique in 13 TeV pp collisions with the ATLAS detector

The ATLAS Collaboration

A measurement of off-shell Higgs boson production in the $H^* \to ZZ \to 4\ell$ decay channel is presented. The measurement uses 140 fb^{-1} of proton–proton collisions at $\sqrt{s} = 13$ TeV collected by the ATLAS detector at the Large Hadron Collider and supersedes the previous result in this decay channel using the same dataset. The data analysis is performed using a neural simulation-based inference method, which builds per-event likelihood ratios using neural networks. The observed (expected) off-shell Higgs boson production signal strength in the $ZZ \to 4\ell$ decay channel at 68% CL is $0.87^{+0.75}_{-0.54}$ ($1.00^{+1.04}_{-0.95}$). The evidence for off-shell Higgs boson production using the $ZZ \to 4\ell$ decay channel has an observed (expected) significance of 2.5σ (1.3σ). The expected result represents a significant improvement relative to that of the previous analysis of the same dataset, which obtained an expected significance of 0.5σ. When combined with the most recent ATLAS measurement in the $ZZ \to 2\ell 2\nu$ decay channel, the evidence for off-shell Higgs boson production has an observed (expected) significance of 3.7σ (2.4σ). The off-shell measurements are combined with the measurement of on-shell Higgs boson production to obtain constraints on the Higgs boson total width. The observed (expected) value of the Higgs boson width at 68% CL is $4.3^{+2.7}_{-1.9}$ ($4.1^{+3.5}_{-3.4}$) MeV.

16:00	Fair Universe HiggsML Uncertainty Challenge	RAGANSU CHAKKAPPAI	Salle Estaque, Palais du Pharo	16:00 - 16:20
	Unbinned machine-learned measurements for the LHC with systematic uncertainties	Robert Schoefbeck	Salle Estaque, Palais du Pharo	16:20 - 16:40
	Higgs Signal Strength Estimation with a Dual-Branch GNN under Systematic Uncertainties	Daohan Wang	Salle Estaque, Palais du Pharo	16:40 - 17:00
17:00	Parameter Estimation with Neural Simulation-Based Inference in ATLAS	Jay Ajitbhai Sandesara	Salle Estaque, Palais du Pharo	17:00 - 17:20
	Constraining the Higgs trilinear self-coupling from off-shell production using neural simulation-based inference	Tae Hyoun Park		
	Multi-Scale Transformer Encoder for Di-Tau Invariant Mass Reconstruction at CMS	Valentina Camagni	Salle Estaque, Palais du Pharo	17:40 - 18:00

simulation-based-inference.org

Simulation-based Inference   Home   Papers   Software   Jobs   Misc   About

Papers

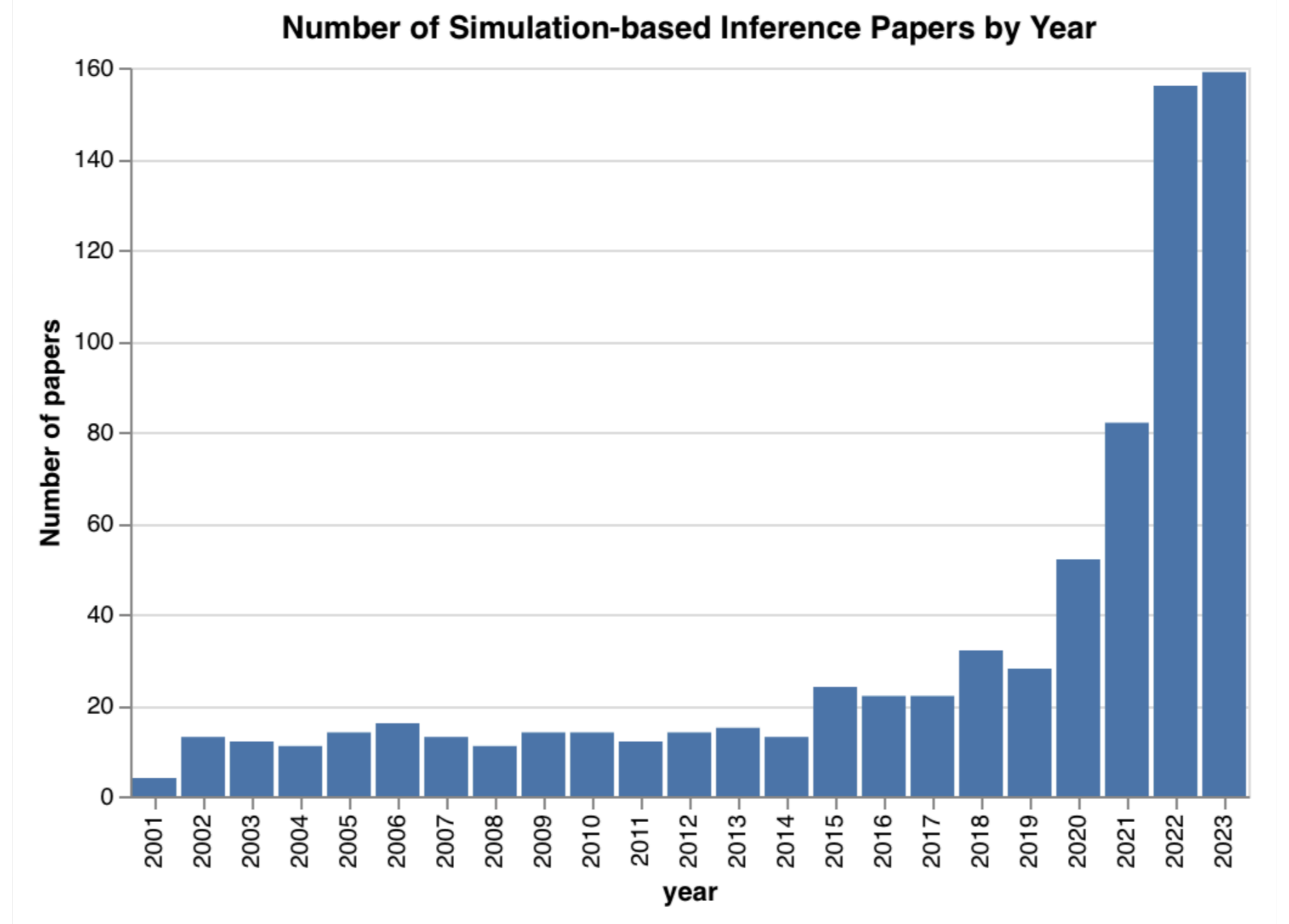
The list is automatically compiled each day. Should you observe any inaccuracies or concerns, kindly bring them to our attention.
Additionally, if you believe a new paper aligns with the topic, feel free to submit it.
Visualize the annual growth in the number of publications.

Sort by Category | Sort by Year | Sort by Journal

- Total (744)
- Statistics (195)
- Computer Science (102)
- Astrophysics (71)
- Mathematics (54)
- Education (47)
- Economics (46)
- Physics (33)
- Quantitative Biology (31)
- Neuroscience (27)
- Quantitative Finance (21)
- Astronomy (14)
- Genetics (13)
- Epidemiology (11)
- Engineering (10)
- Medicine (8)
- Geography (8)
- Social Science (7)
- Evolutionary biology (6)
- Ecology (5)
- Cognitive Science (4)
- Robotics (4)
- Systems biology (4)

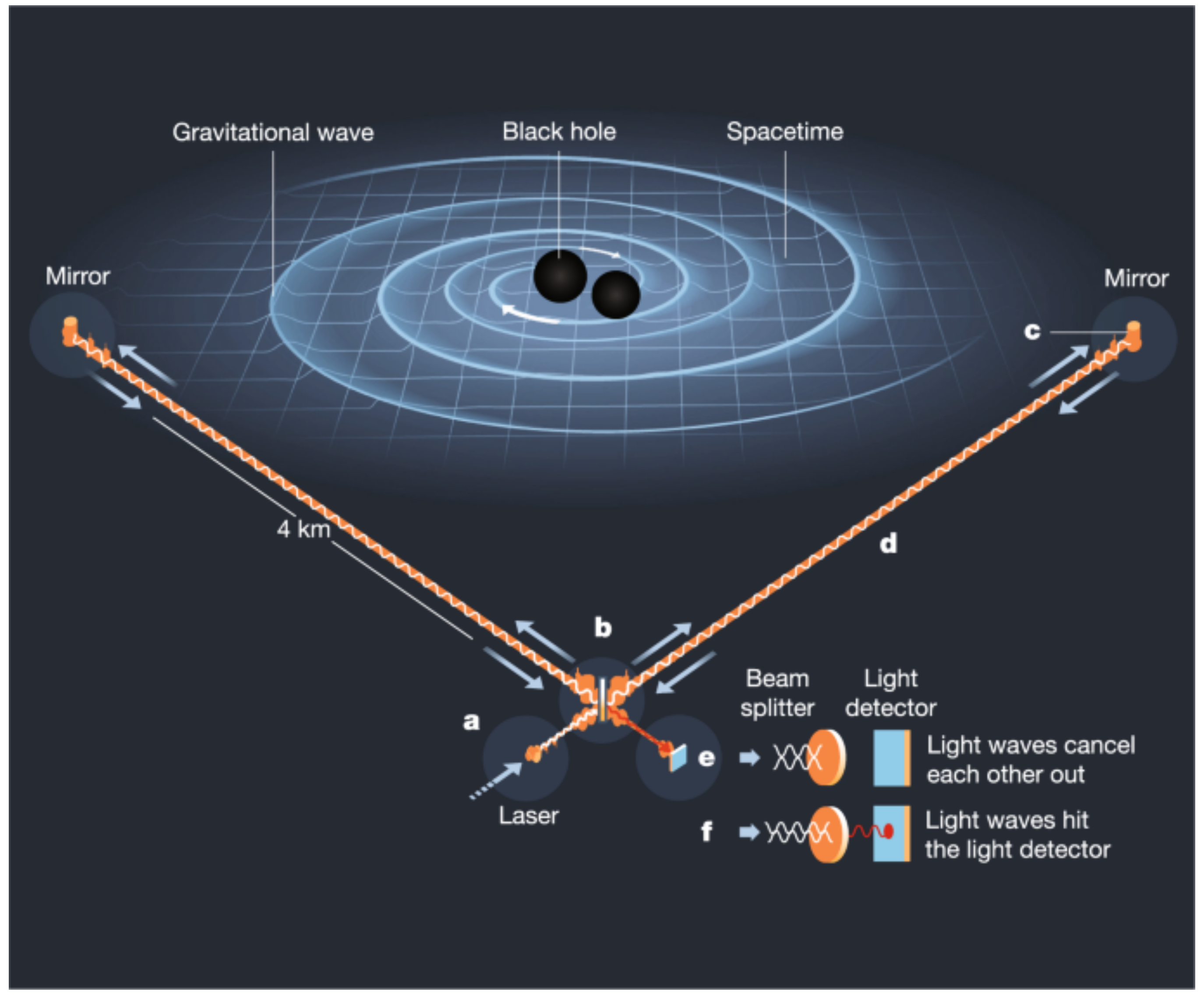
Statistics

- Simulation based stacking, Y Yao, BRS Blancard, J Domke - arXiv preprint arXiv:2310.17009, 2023 - arxiv.org
- Calibrating Neural Simulation-Based Inference with Differentiable Coverage Probability, M Falkiewicz, N Takeishi, I Shekhzadeh... - arXiv preprint arXiv ..., 2023 - arxiv.org
- Simulation-based Inference with the Generalized Kullback-Leibler Divergence, BK Miller, M Federici, C Weniger, P Forré - arXiv preprint arXiv ..., 2023 - arxiv.org
- Simulation-based Inference for Cardiovascular Models, A Wehenkel, J Behrmann, AC Miller, G Sapiro... - arXiv preprint arXiv ..., 2023 - arxiv.org
- Hierarchical Neural Simulation-Based Inference Over Event Ensembles, L Heinrich, S Mishra-Sharma, C Pollard... - arXiv preprint arXiv ..., 2023 - arxiv.org
- L-C2ST Local Diagnostics for Posterior Approximations in Simulation-Based Inference, J Linhart, A Gramfort, PLC Rodrigues - arXiv preprint arXiv:2306.03580, 2023 - arxiv.org
- Learning Robust Statistics for Simulation-based Inference under Model Misspecification, D Huang, A Bharti, A Souza, L Acerbi... - arXiv preprint arXiv ..., 2023 - arxiv.org
- Generalized Bayesian Inference for Scientific Simulators via Amortized Cost Estimation, R Gao, M Deistler, JH Macke - arXiv preprint arXiv:2305.15208, 2023 - arxiv.org
- Variational Inference with Coverage Guarantees, Y Patel, D McNamara, J Loper, J Regier... - arXiv preprint arXiv ..., 2023 - arxiv.org
- Generalised likelihood profiles for models with intractable likelihoods, DJ Warne, OJ Maclaren, EJ Carr, MJ Simpson... - arXiv preprint arXiv ..., 2023 - arxiv.org
- Neural Likelihood Surfaces for Spatial Processes with Computationally Intensive or Intractable Likelihoods, J Walchessen, A Lenzi, M Kuusela - arXiv preprint arXiv:2305.04634, 2023 - arxiv.org
- Balancing Simulation-based Inference for Conservative Posteriors, A Delaunoy, BK Miller, P Forré, C Weniger... - arXiv



See also: github.com/smsharma/awesome-neural-sbi

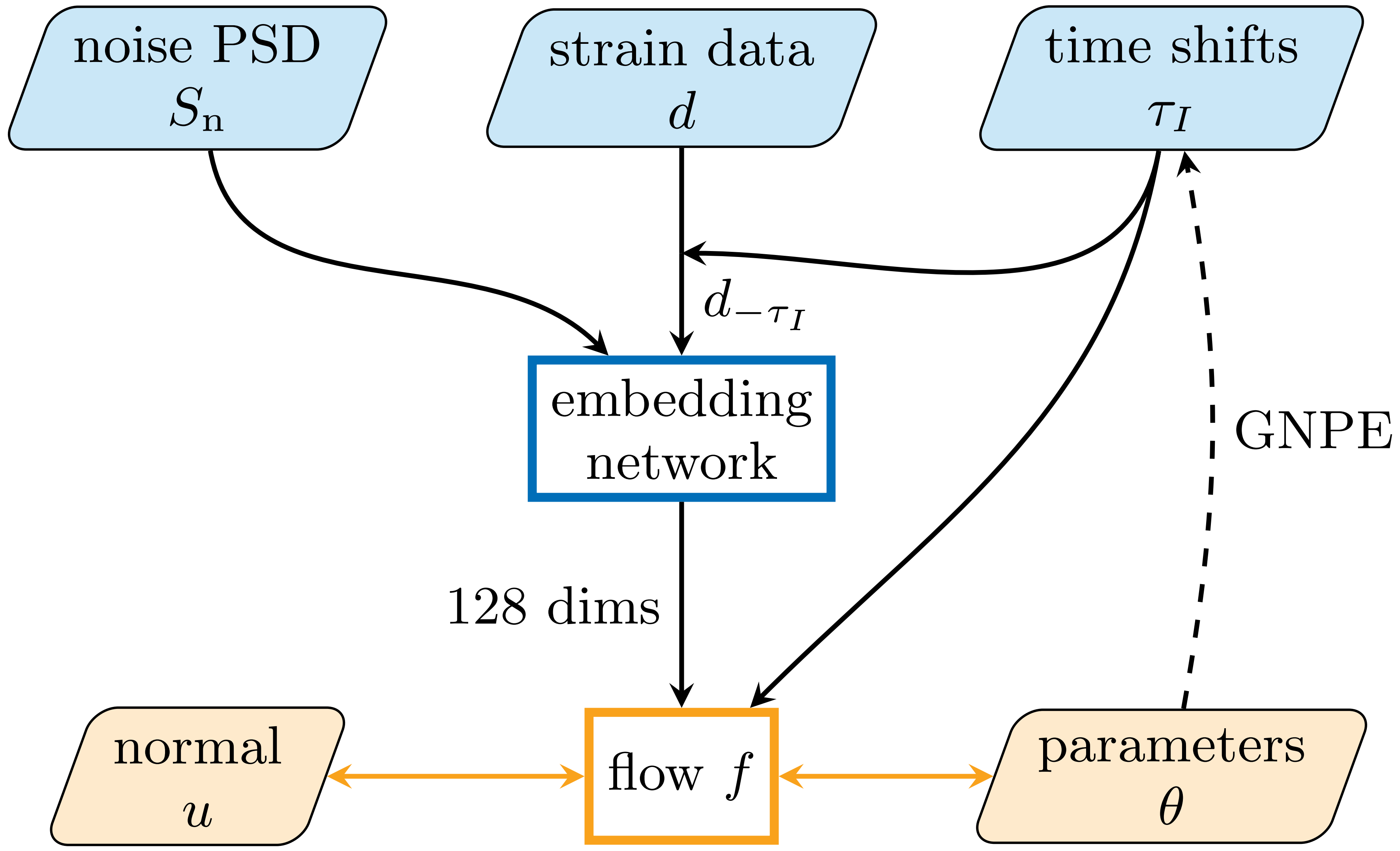
Gravitational Wave Astronomy



Real-time gravitational-wave science with neural posterior estimation

Maximilian Dax,[1,*] Stephen R. Green,[2,†] Jonathan Gair,[2,‡]
Jakob H. Macke,[1,3] Alessandra Buonanno,[2,4] and Bernhard Schölkopf[1]

[1] *Max Planck Institute for Intelligent Systems, Max-Planck-Ring 4, 72076 Tübingen, Germany*
[2] *Max Planck Institute for Gravitational Physics (Albert Einstein Institute), Am Mühlenberg 1, 14476 Potsdam, Germany*
[3] *Machine Learning in Science, University of Tübingen, 72076 Tübingen, Germany*
[4] *Department of Physics, University of Maryland, College Park, MD 20742, USA*



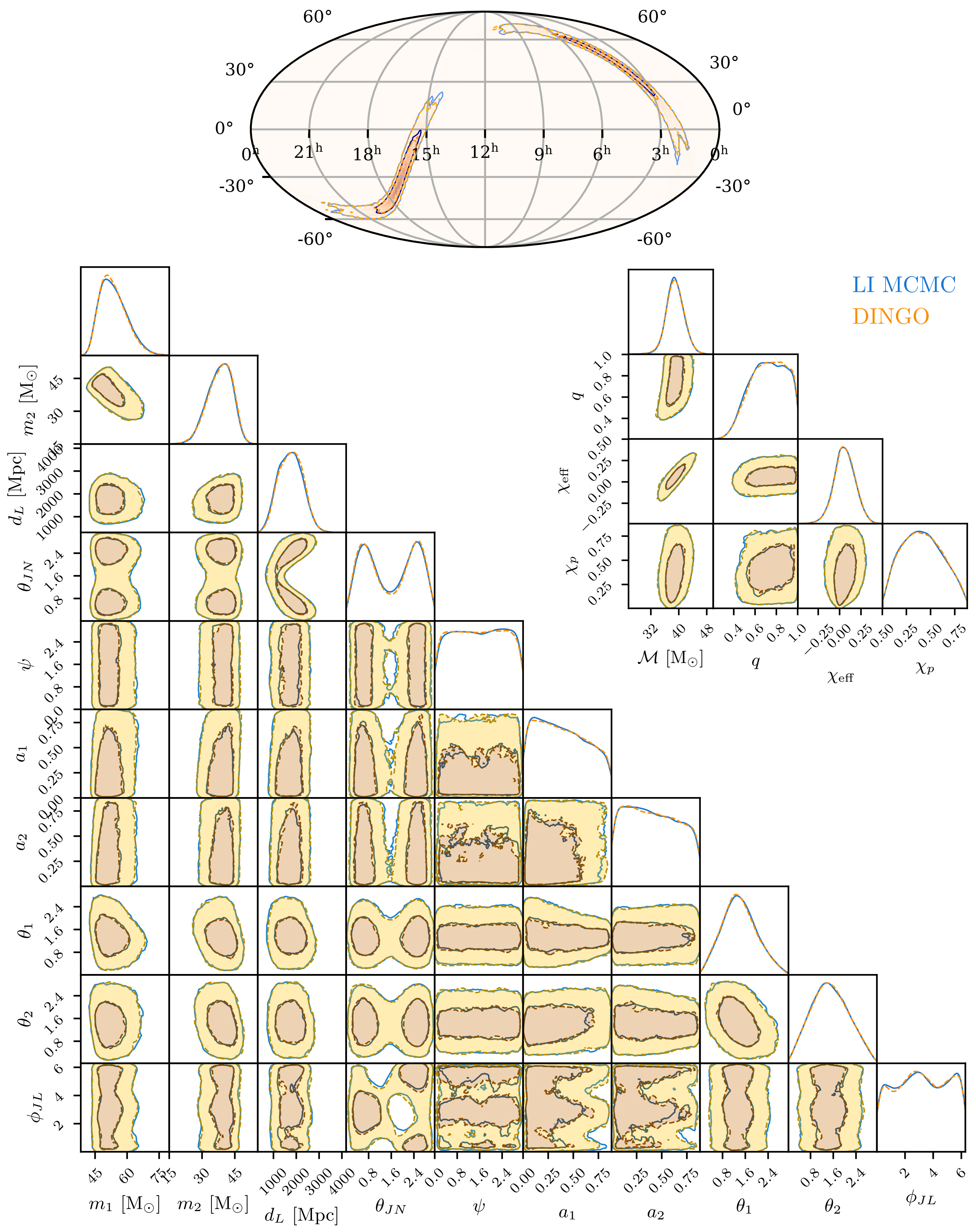


Figure 15. GW170823.

We are at a tipping point for SBI in HEP

Dedicated discussion at EuCAIFCon2025 — a new working group?

Discussion: Simulation-based inference & Uncertainty quantification

- Simulation-based inference plays an increasing role in cosmology, gravitational waves, astroparticle physics and particle/nuclear physics
- Still there are lots of practical hurdles to make SBI a standard workhorse for analysis tasks.

- Goal of the discussion is to establish the most critical needs in the community (common tools, large joined projects, training material, algorithmic gaps, etc).
- Identify clear goals and timelines to form a workgroup

Similar to the transition we made to statistical procedures at the LHC used for Higgs discovery

- A forum for discussions
- Shared formalism, conventions, & recommendations
- Benchmark examples to build trust
- Guidance and prioritization for tool developers

2024: https://indico.nikhef.nl/event/4875/
2025: https://agenda.infn.it/event/43565/

Computational Microscope

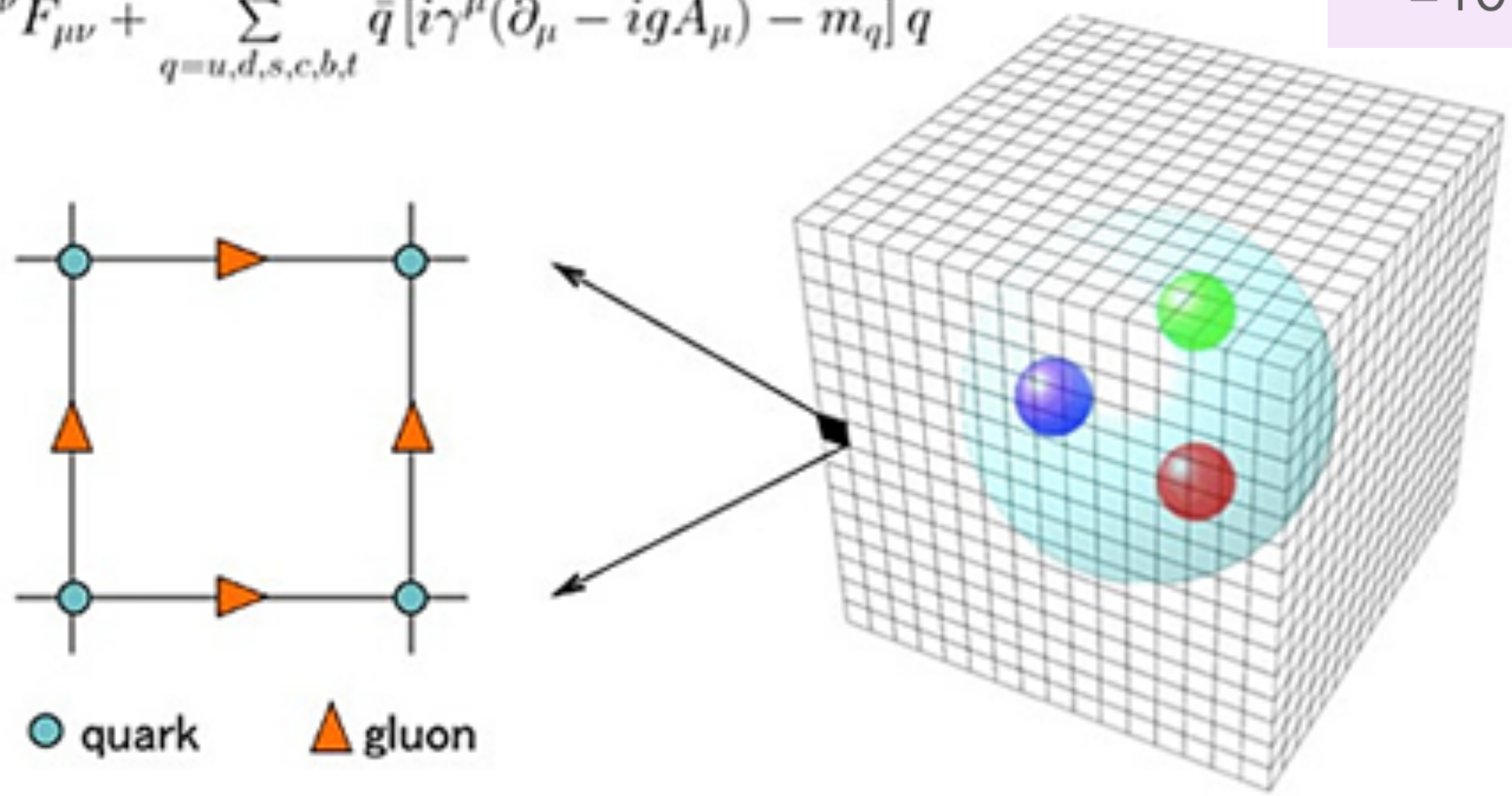
Lattice Field Theory

Lattice field theory is a computational approach to studying field theory on a discretized space-time.

- **Path integral**: a "path" is a sample from distribution of lattice configurations **~exp(-Action[path])**
- Predictions are **expectations** of quantum operators w.r.t. this distribution.
- **Hamiltonian Monte Carlo was invented for this problem**, but it has limitations.

QCD Lagrangian

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \sum_{q=u,d,s,c,b,t} \bar{q}\left[i\gamma^{\mu}(\partial_{\mu} - igA_{\mu}) - m_q\right]q$$



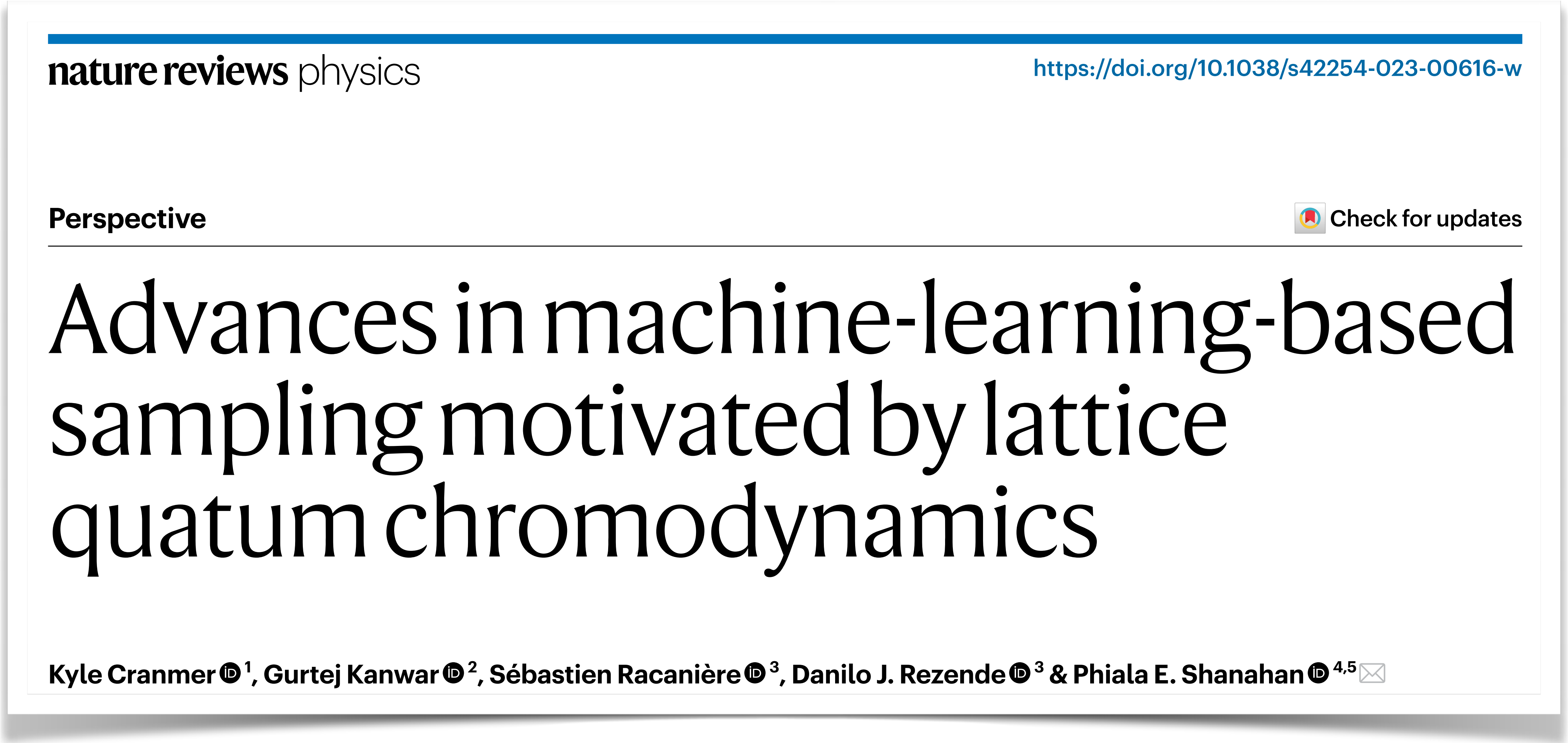
$64^3 \times 128 \times 4 \times 9 \times 2$
$\simeq 10^9$ numbers

AI-Enhanced Monte Carlo Integration

Basic idea:

- use **generative AI** model (normalizing flows) to approximate the target Boltzmann distribution.
- Sample from the generative AI model instead of traditional Hamiltonian MC

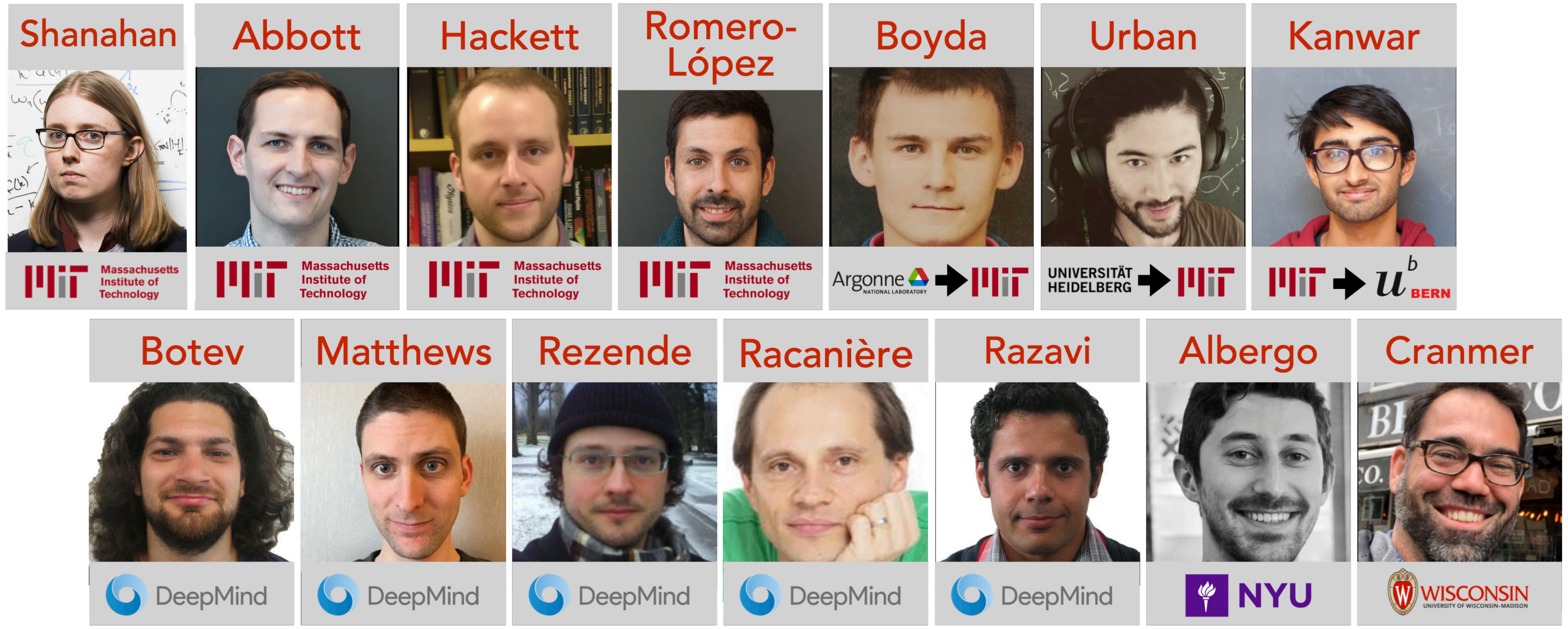
Learned model won't be perfect, but you can **correct** via importance sampling or MCMC procedure


nature reviews physics https://doi.org/10.1038/s42254-023-00616-w

Perspective Check for updates

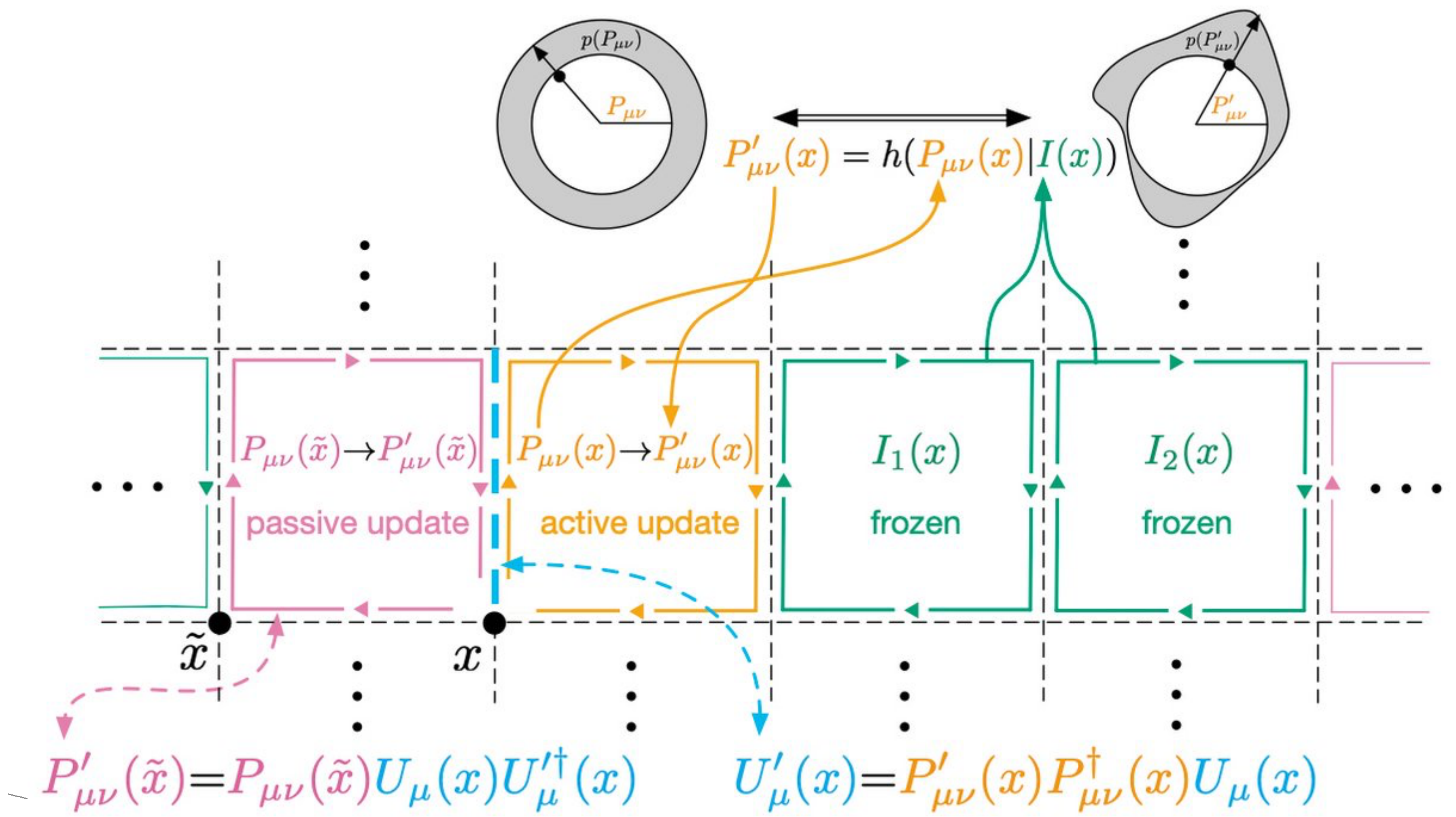
Advances in machine-learning-based sampling motivated by lattice quatum chromodynamics

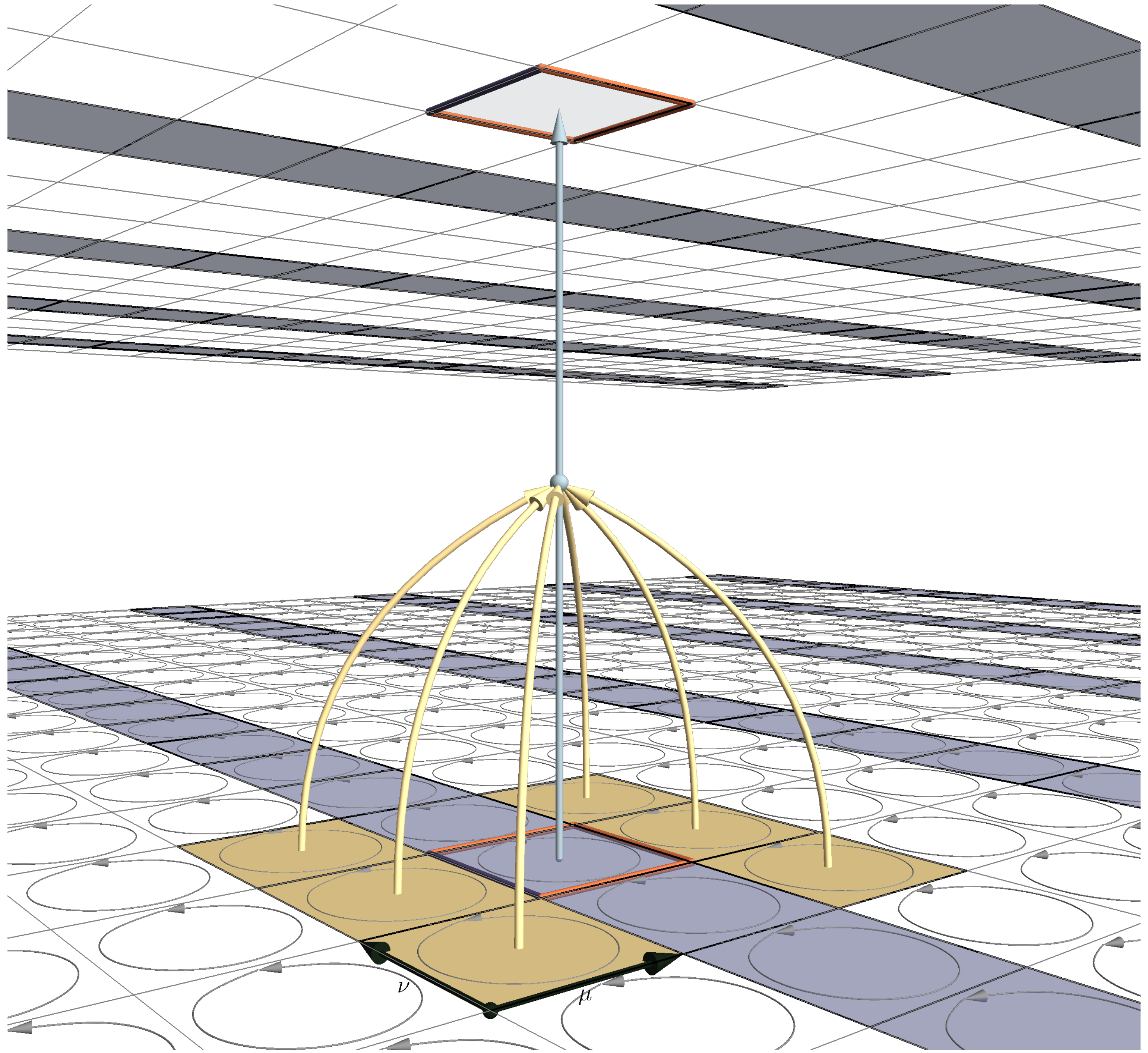
Kyle Cranmer[1], Gurtej Kanwar[2], Sébastien Racanière[3], Danilo J. Rezende[3] & Phiala E. Shanahan[4,5]

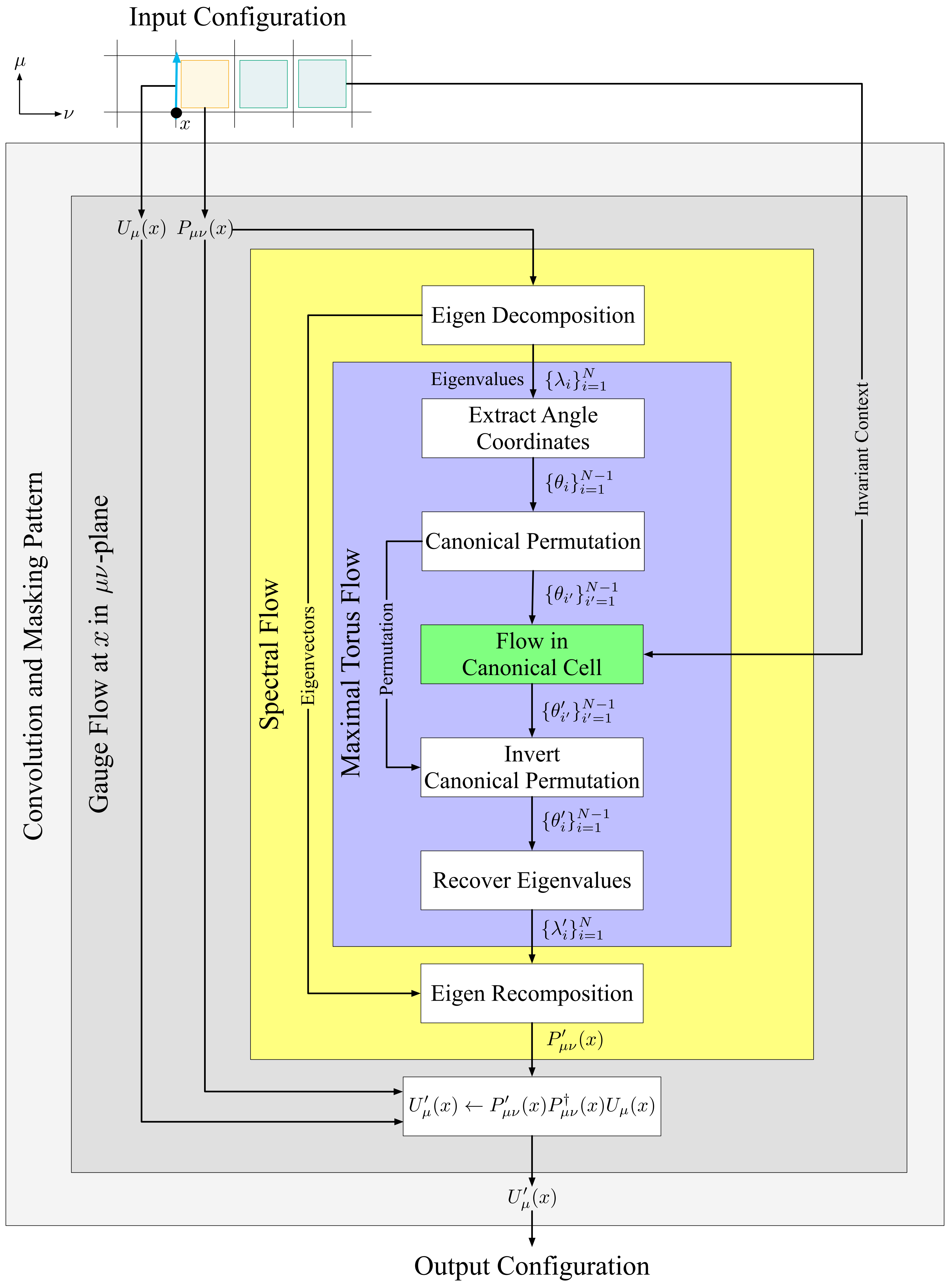


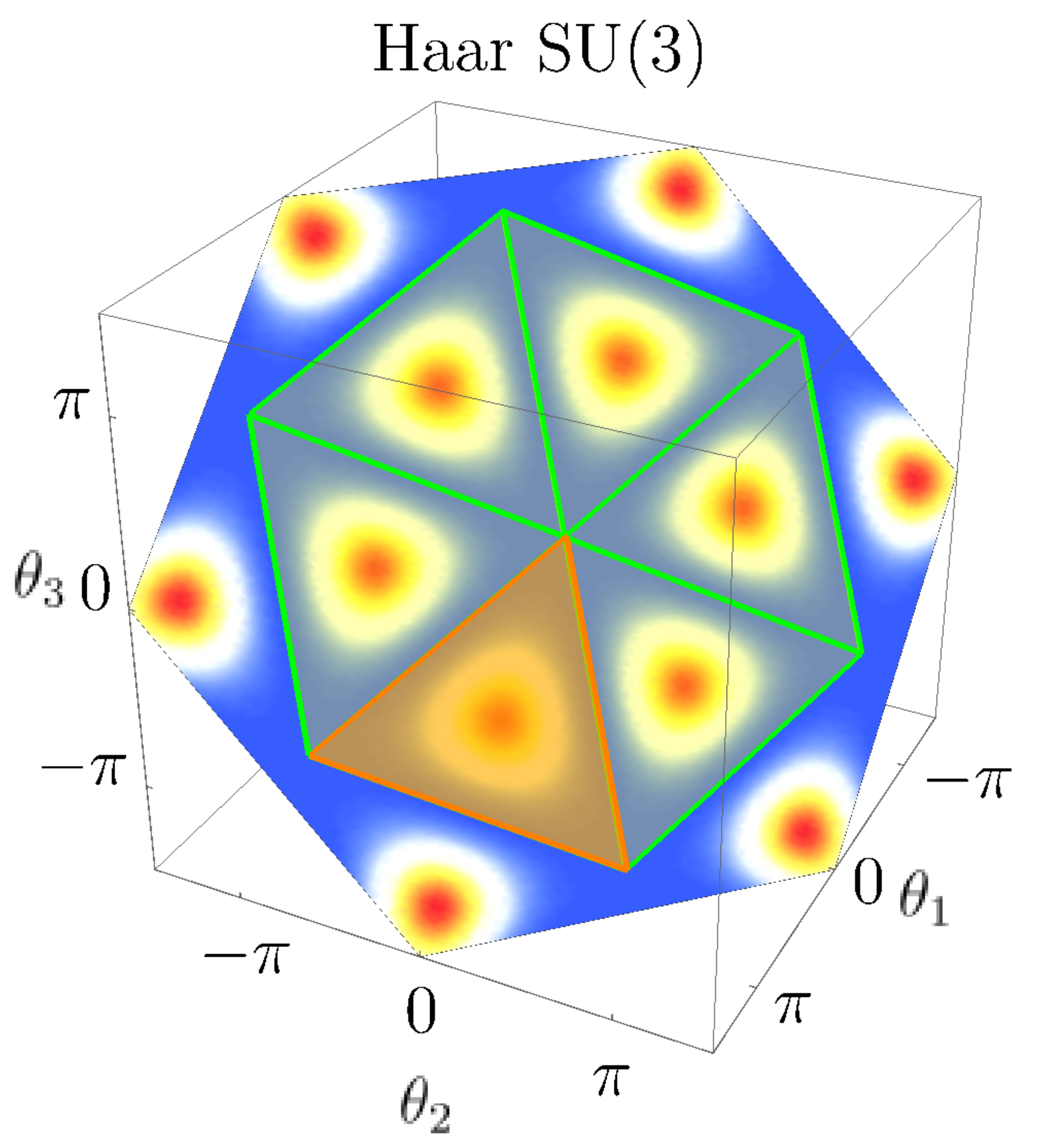
See also: Albergo, Kanwar, Shanahan, PRD (2019) arXiv:1904.12072

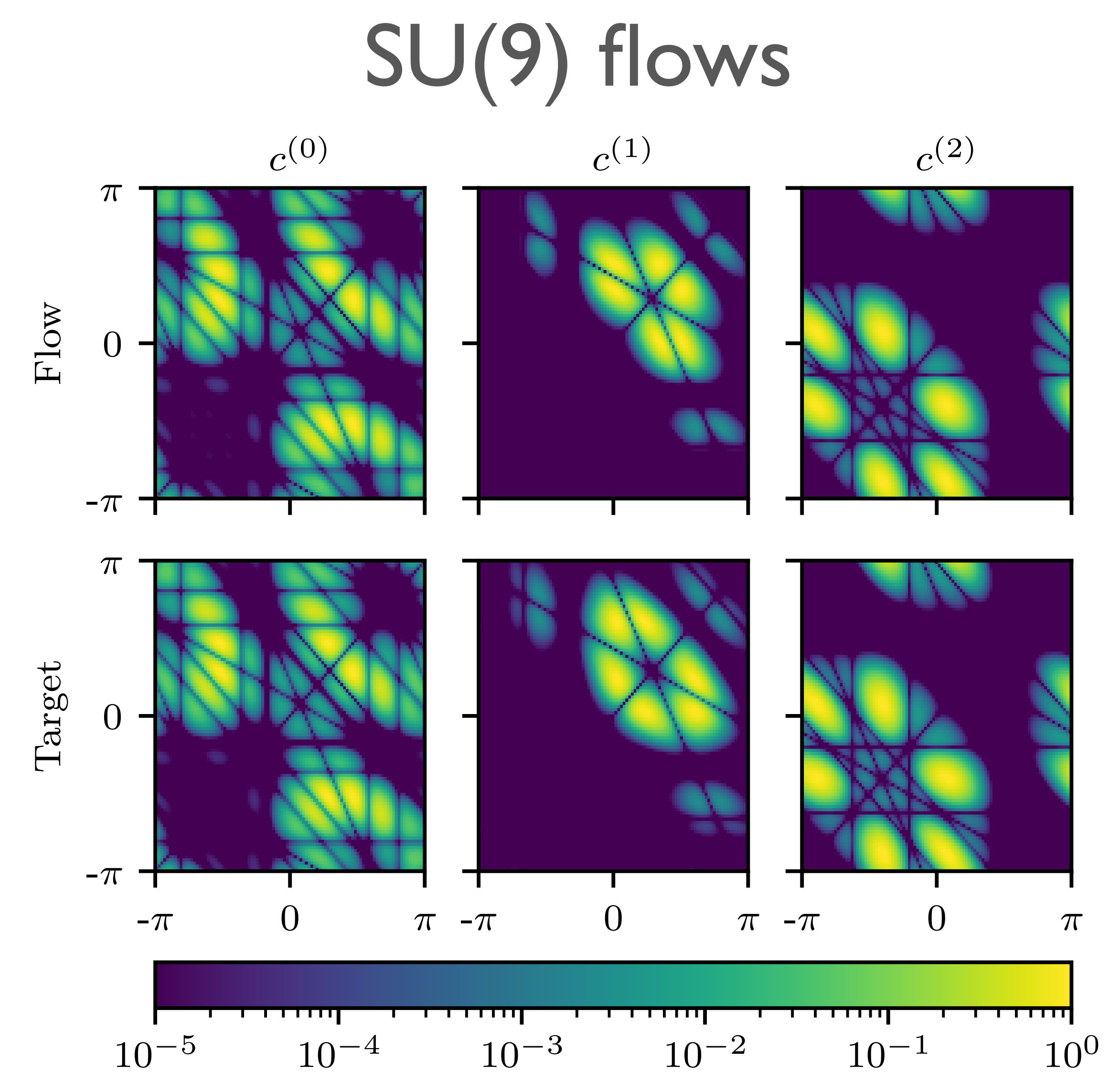
Space-time & Local, Non-Abelian Gauge Symmetry









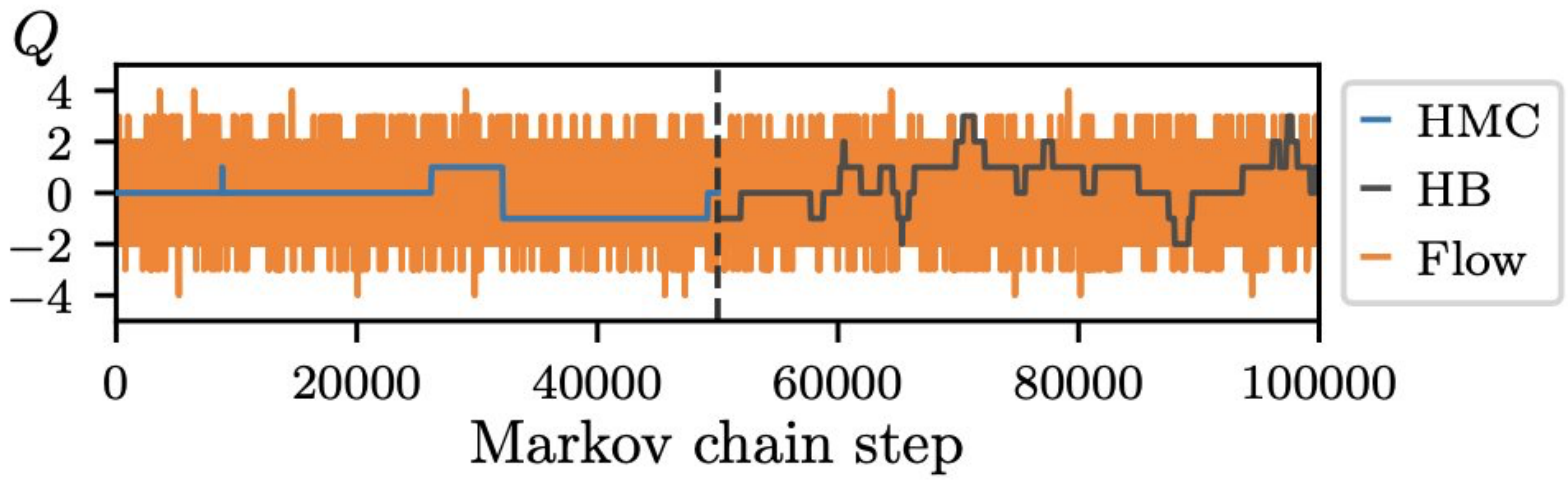


A promising direction

Phys.Rev.Lett. 125 (2020) 12, 121601 [arXiv:2003.06413]

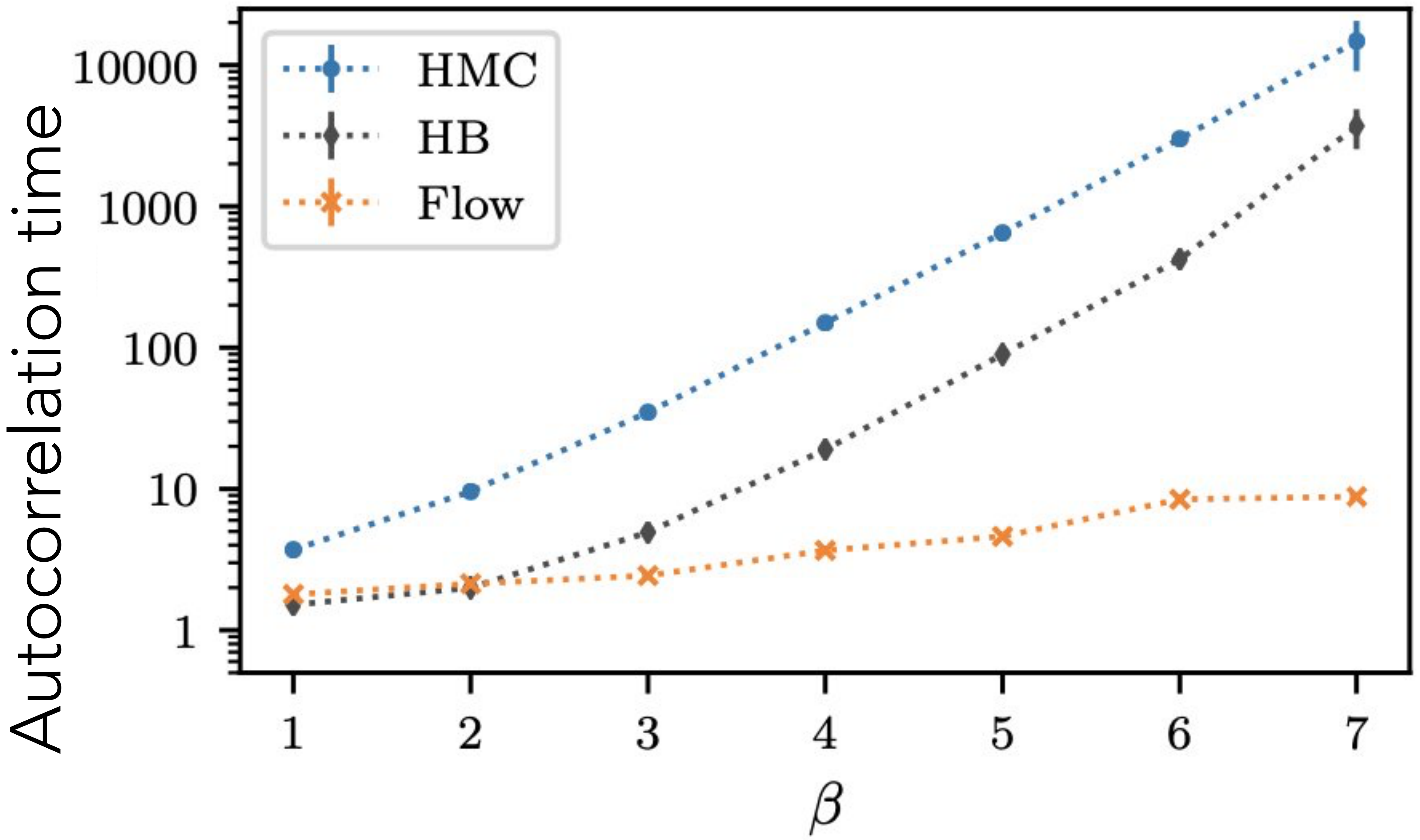
Essentially, MCMC can get stuck for a while in a certain mode.

- Our new “flow-based” proposal does much better!
- It learns to propose configurations that look like our target distribution.
- 1000x reduction in autocorrelation time



The topological charge Q will be constant for thousands of MCMC steps.

2D U(1) model



For molecular dynamics

RESEARCH Noé *et al.*, *Science* **365**, 1001 (2019) 6 September 2019

RESEARCH ARTICLE SUMMARY

MACHINE LEARNING

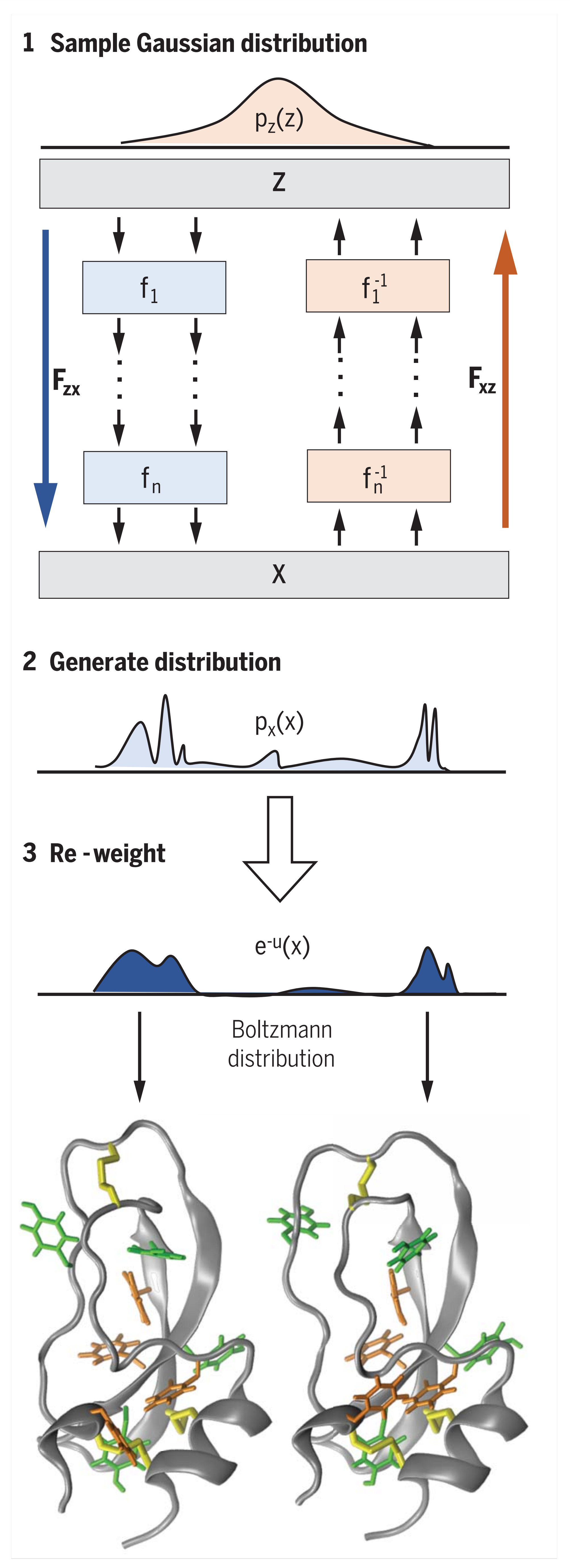
Boltzmann generators: Sampling equilibrium states of many-body systems with deep learning

Frank Noé*†, Simon Olsson*, Jonas Köhler*, Hao Wu

The main approach is thus to start with one configuration, e.g., the folded protein state, and make tiny changes to it over time, e.g., by using Markov-chain Monte Carlo or molecular dynamics (MD). However, these simulations get trapped in metastable (long-lived) states: For example, sampling a single folding or unfolding event with atomistic MD may take a year on a supercomputer.

Same core idea as Boltzmann generators

Boltzmann generators overcome sampling problems between long-lived states.



Similar ideas are improving phase space integration to accelerate Parton-level Monte Carlo generators

MadNIS – Neural Multi-Channel Importance Sampling

Theo Heimel[1], Ramon Winterhalder[2],
Anja Butter[1,3], Joshua Isaacson[4], Claudius Krause[1],
Fabio Maltoni[2,5], Olivier Mattelaer[2], and Tilman Plehn[1]

1 Institut für Theoretische Physik, Universität Heidelberg, Germany
2 CP3, Université catholique de Louvain, Louvain-la-Neuve, Belgium
3 LPNHE, Sorbonne Université, Université Paris Cité, CNRS/IN2P3, Paris, France
4 Theoretical Physics Division, Fermi National Accelerator Laboratory, Batavia, IL, USA
5 Dipartimento di Fisica e Astronomia, Universitá di Bologna, Italy

ramon.winterhalder@uclouvain.be

Abstract

Theory predictions for the LHC require precise numerical phase-space integration and generation of unweighted events. We combine machine-learned multi-channel weights with a normalizing flow for importance sampling, to improve classical methods for numerical integration. We develop an efficient bi-directional setup based on an invertible network, combining online and buffered training for potentially expensive integrands. We illustrate our method for the Drell-Yan process with an additional narrow resonance.

3.2 Neural importance sampling

Second, MADNIS augments the physics-inspired phase space mappings with an INN [19]

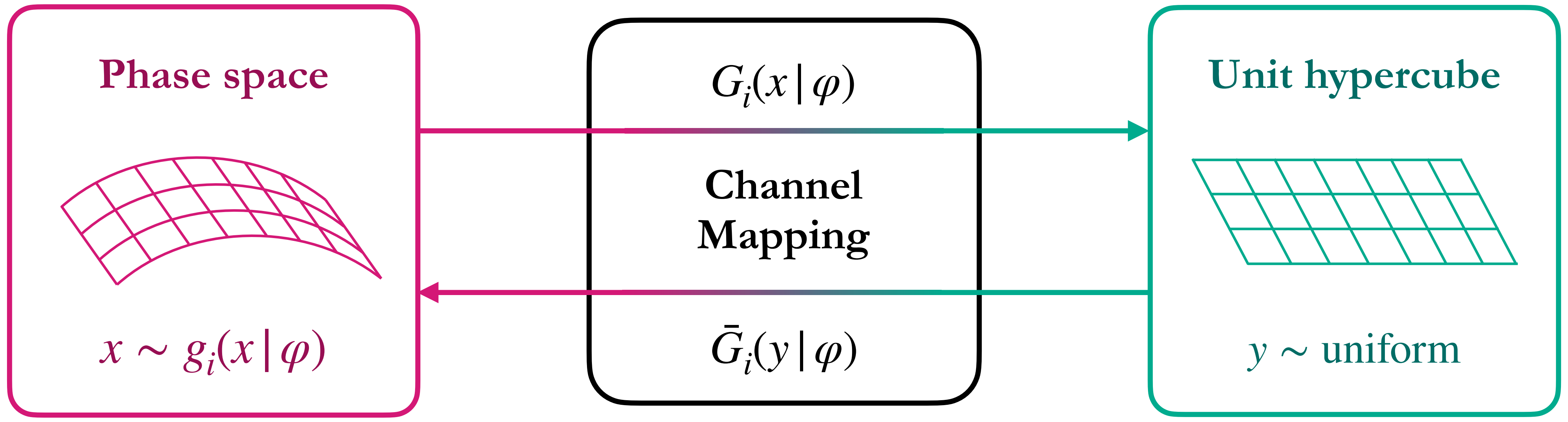
$$y = G_i(x) \to G_i(x|\varphi) \qquad \text{and} \qquad x = \overline{G}_i(y|\varphi)\,. \tag{22}$$

This replaces the classic importance sampling density $g_i(x)$ with a network-based variable transformation $g_i(x|\varphi)$ in Eqs.(6) and (15)

$$I[f] = \sum_i \int_{U_i} \mathrm{d}^d y\ \alpha_i(x) \frac{f(x)}{g_i(x|\varphi)}\bigg|_{x=\overline{G}_i(y|\varphi)} \quad \text{with} \quad g_i(x|\varphi) = \left|\frac{\partial G_i(x|\varphi)}{\partial x}\right|\,, \tag{23}$$

where we assume the latent distribution in y to be uniform. The INN-encoded phase space mapping is trained to provide a surrogate density

$$g_i(x|\varphi) \approx f_i(x) = \alpha_i(x) f(x)\,, \tag{24}$$



Conclusion

AI/ML is providing qualitatively new capabilities.

- Those capabilities allow us to remove some approximations & simplifications and return to what we've always wanted to do from first principles but were unable to do computationally.

It is allowing us to make better use of our experimental data and theoretical understanding.

- Many challenges remain, the transformation is far from over.

Backup

Statistical Framing

forward modeling
generation
simulation

PREDICTION

$p(\,x, z \mid \theta\,)$

$\boldsymbol{\theta}$
parameters of interest

$\mathbf{z}$
latent variables

$\mathbf{x}$
observed data
simulated data

INFERENCE

inverse problem
measurement
parameter estimation

Simulating particle physics processes

Theory
parameters
θ

Evolution

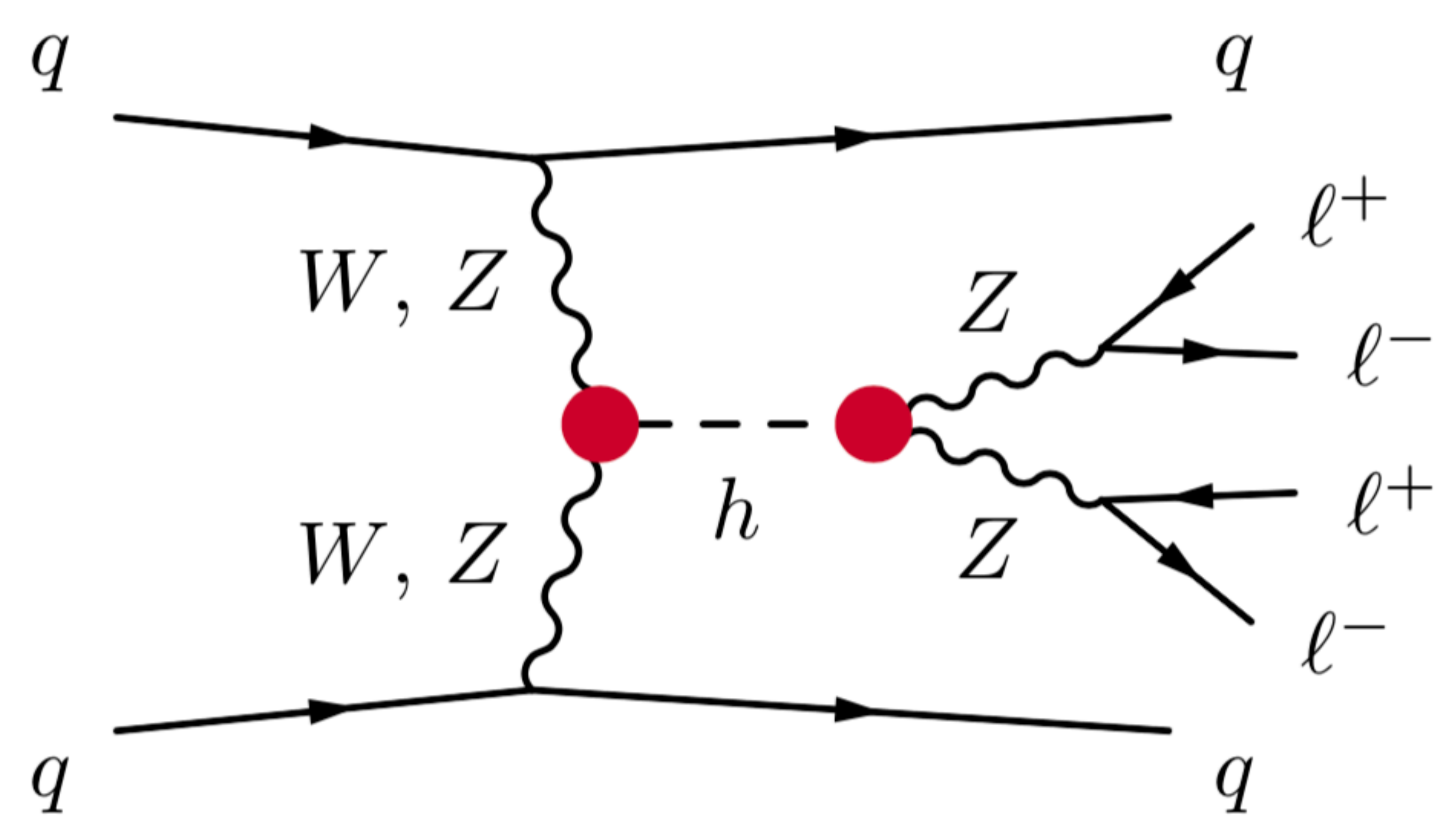
Simulating particle physics processes

Latent variables

Parton-level momenta

Theory parameters

$z_p \longleftarrow \theta$



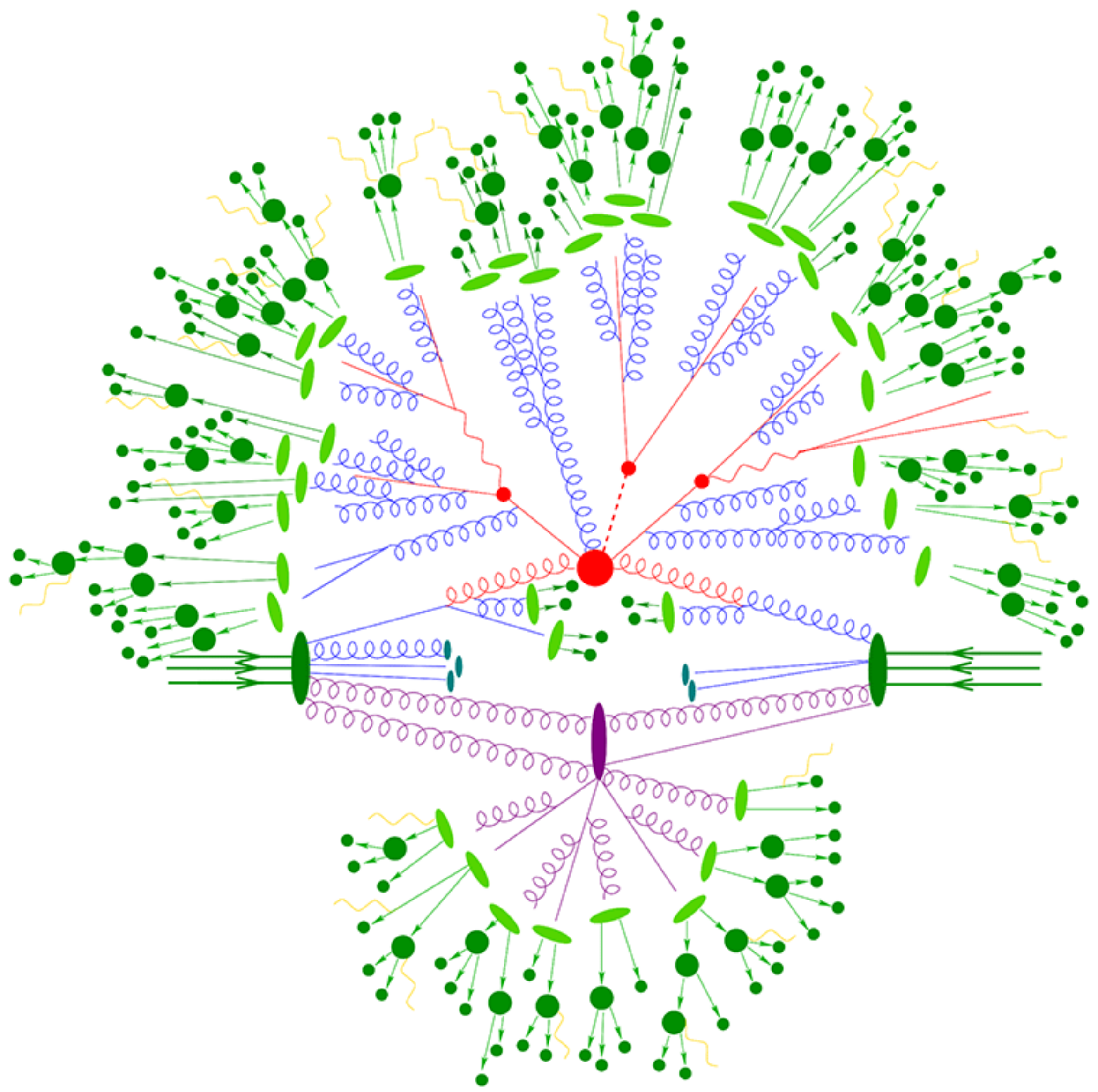
Evolution

Simulating particle physics processes

Latent variables

Shower splittings | Parton-level momenta | Theory parameters

$z_s \longleftarrow z_p \longleftarrow \theta$



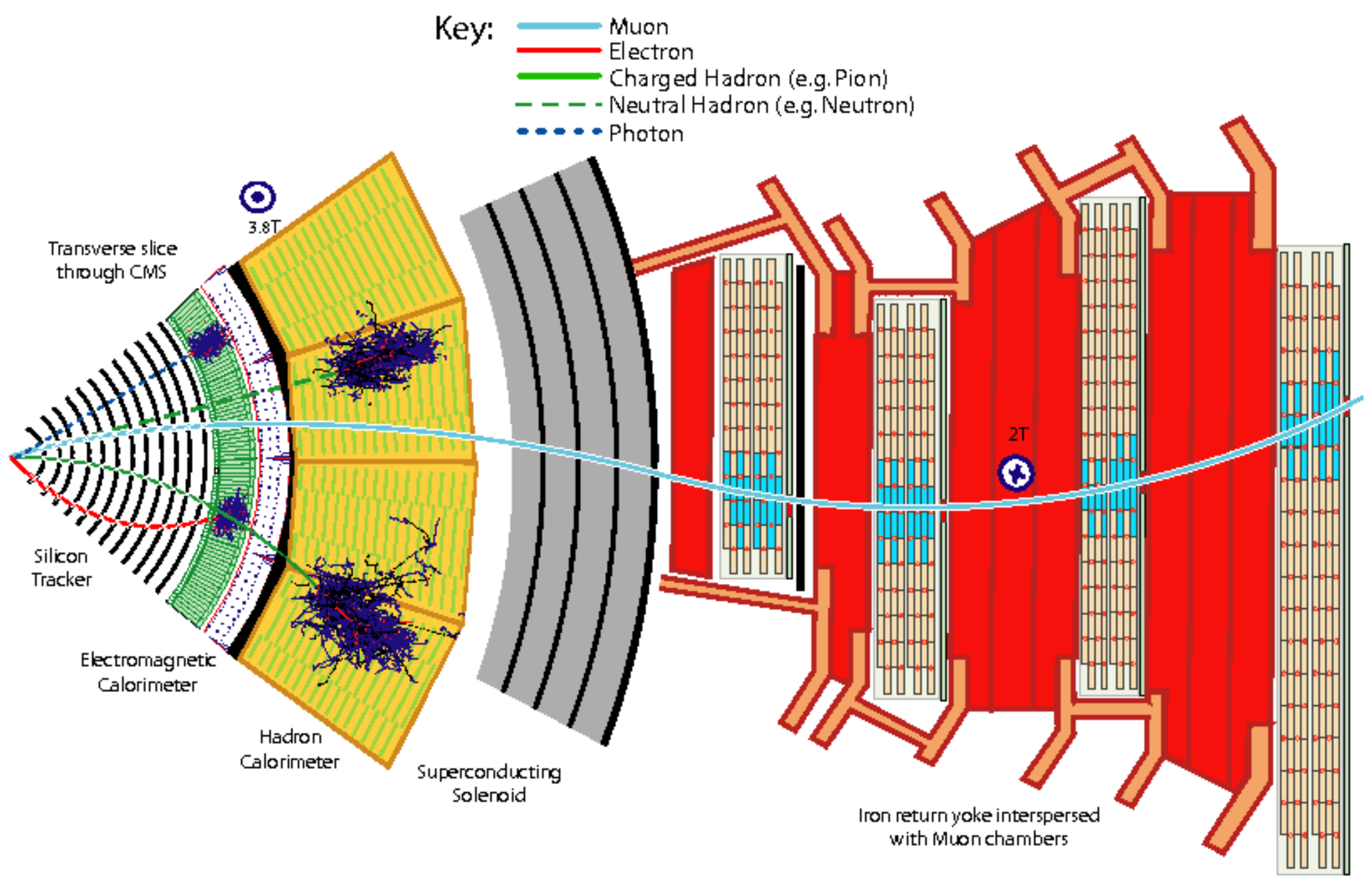
[F. Krauss]

Evolution

Simulating particle physics processes

Latent variables

Detector interactions	Shower splittings	Parton-level momenta	Theory parameters
z_d ←	z_s ←	z_p ←	θ



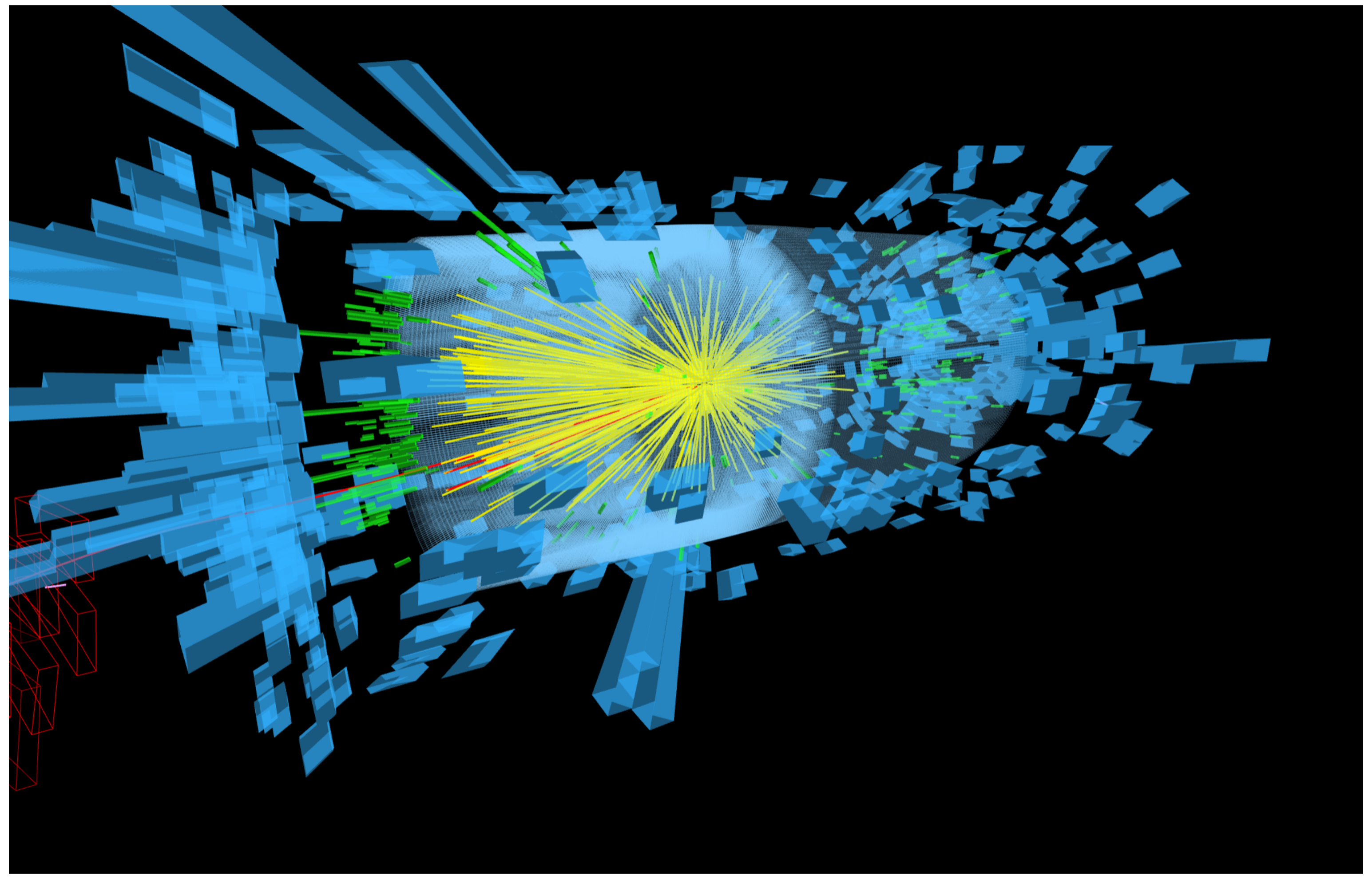
[CMS]

Evolution

Simulating particle physics processes

Latent variables

Observables		Detector interactions		Shower splittings		Parton-level momenta		Theory parameters
x	$\leftarrow$	z_d	$\leftarrow$	z_s	$\leftarrow$	z_p	$\leftarrow$	θ



Evolution

Simulating particle physics processes

Latent variables

Observables		Detector interactions		Shower splittings		Parton-level momenta		Theory parameters
x	$\leftarrow$	z_d	$\leftarrow$	z_s	$\leftarrow$	z_p	$\leftarrow$	θ

Sample from $p(x|z_d)$ $p(z_d|z_s)$ $p(z_s|z_p)$ $p(z_p|\theta)$

GEANT4 A SIMULATION TOOLKIT

MADGRAPH5_aMC@NLO

Prediction (simulation)

Simulating particle physics processes

Latent variables

Observables	Detector interactions	Shower splittings	Parton-level momenta	Theory parameters
x	z_d	z_s	z_p	θ

$$x \longleftarrow z_d \longleftarrow z_s \longleftarrow z_p \longleftarrow \theta$$

$$p(x|\theta) = \int \mathrm{d}z_d \int \mathrm{d}z_s \int \mathrm{d}z_p \quad p(x|z_d) \quad p(z_d|z_s) \quad p(z_s|z_p) \quad p(z_p|\theta)$$

$\longrightarrow$ Inference

Simulating particle physics processes

Latent variables

Observables	Detector interactions	Shower splittings	Parton-level momenta	Theory parameters
x	z_d	z_s	z_p	θ

$$x \longleftarrow z_d \longleftarrow z_s \longleftarrow z_p \longleftarrow \theta$$

$$p(x|\theta) = \boxed{\int \mathrm{d}z_d \int \mathrm{d}z_s \int \mathrm{d}z_p} \; p(x|z_d) \qquad p(z_d|z_s) \qquad p(z_s|z_p) \qquad p(z_p|\theta)$$

It's infeasible to calculate the integral over this enormous space!

$\longrightarrow$

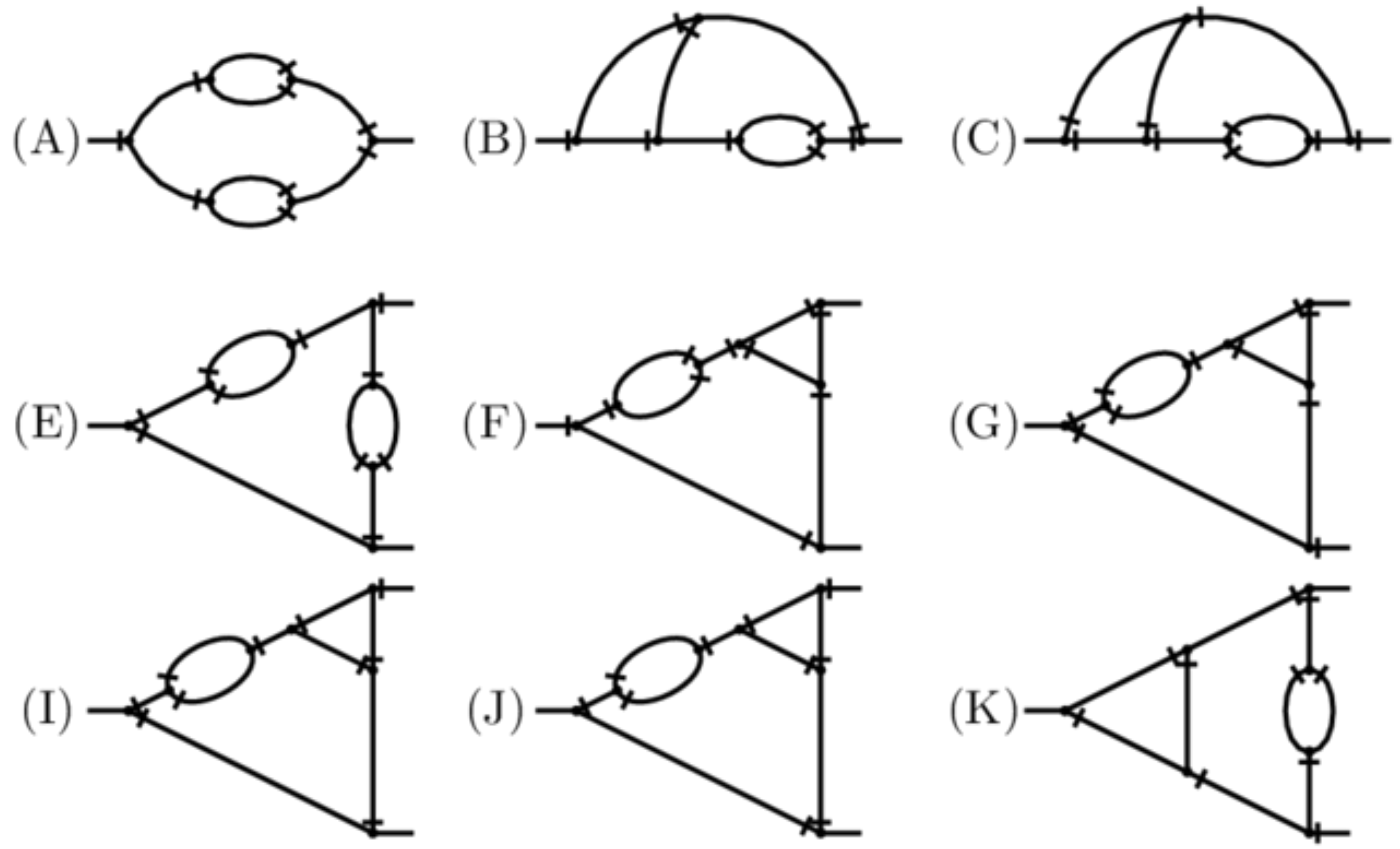
Inference

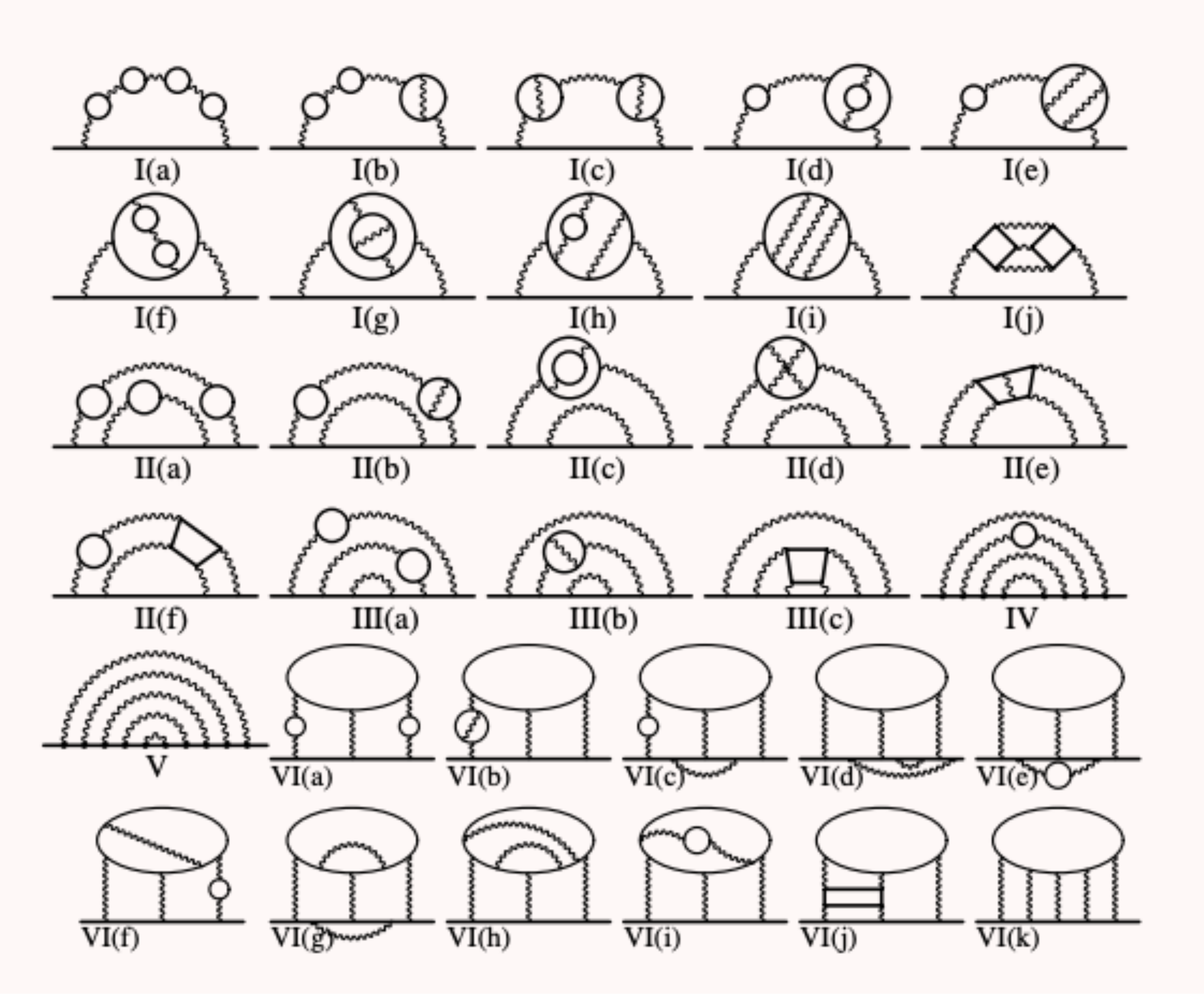
Feynman diagrams with loops

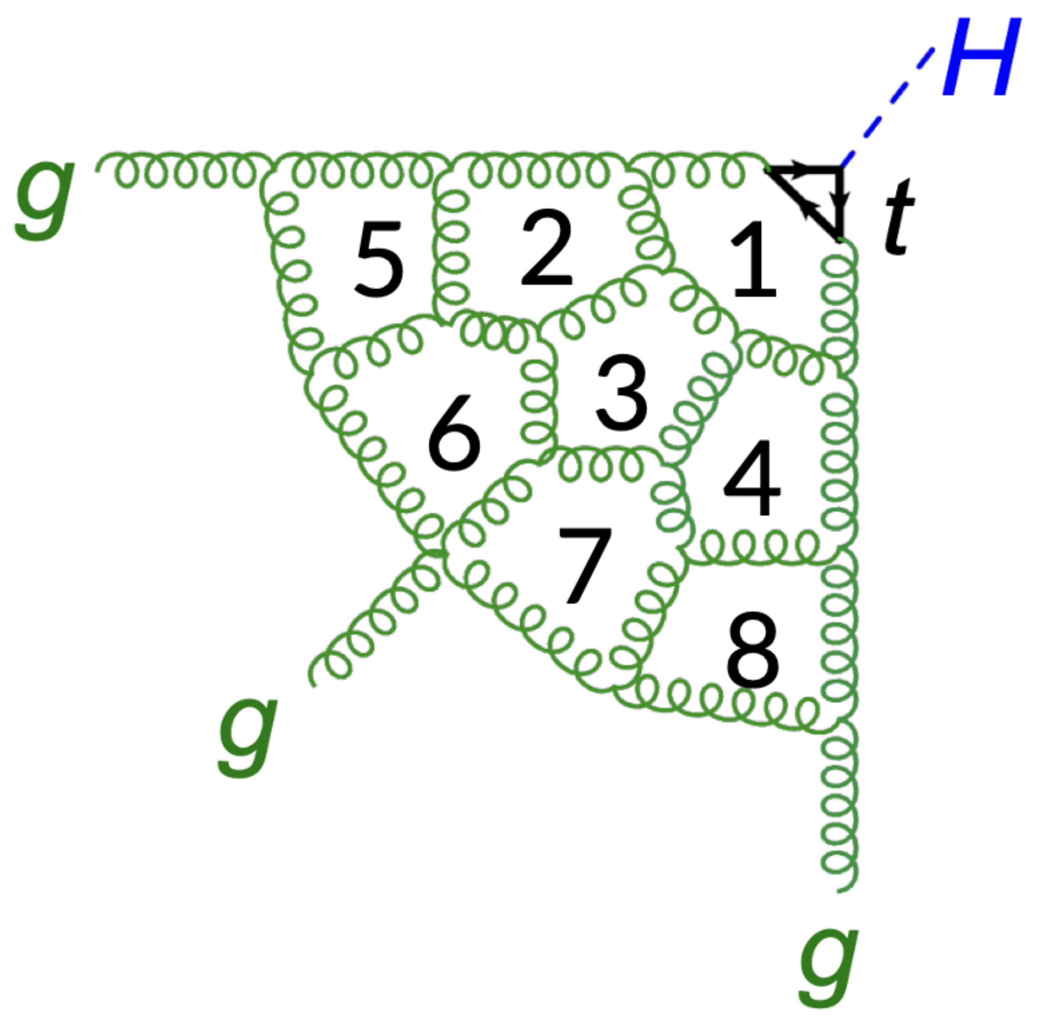
More precise calculations have more **loops**

- But the number of diagrams grows combinatorially with the number of loops
- Feynman diagrams become a poor way to organize the calculation

New **bootstrap** approach emerged that leverages analytical properties of amplitudes. Properties are so constraining, they define a unique solution







In this bootstrap approach the L-loop amplitude can be expressed as a sum of **terms** with an **integer coefficient** and a **word** composed of 2L **letters**

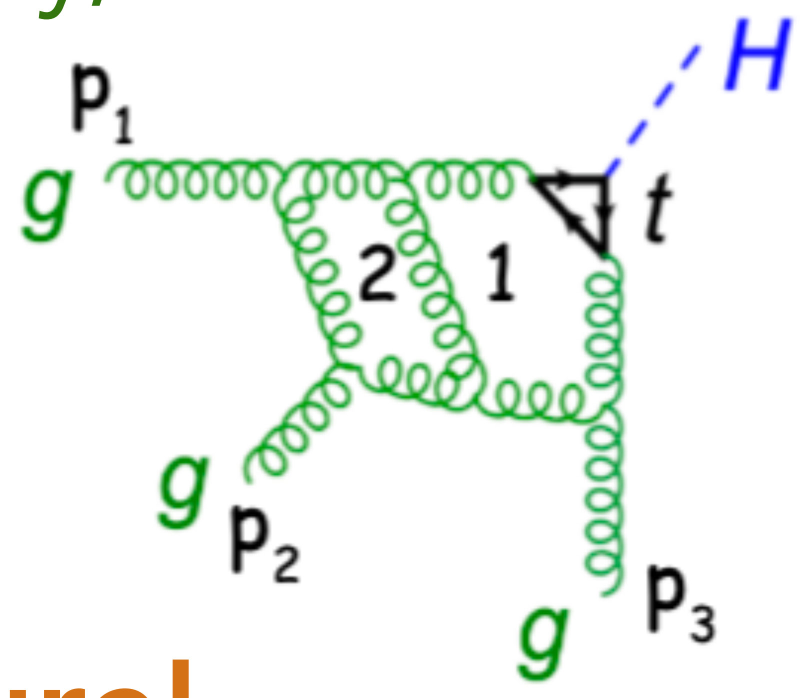
- The 6 letters $\{a, b, c, d, e, f\}$ encode the kinematics of the collision

The Bootstrap

In this bootstrap approach the L-loop amplitude can be expressed as a sum of **terms** with an **integer coefficient** and a **word** composed of 2L **letters**

- The 6 letters $\{a, b, c, d, e, f\}$ encode the kinematics of the collision

For example, in a particular theory called $\mathcal{N}=4$ Super Yang-Mills theory, the answer at 2-loops for a particular interaction is:



$$\mathcal{S}[F_3^{(2)}] = +8bddd +8ceee +8afff +8bfff +8cddd +8aeee$$
$$+16bbbd+16ccce+16bbbf+16aaaf+16cccd+16aaae$$

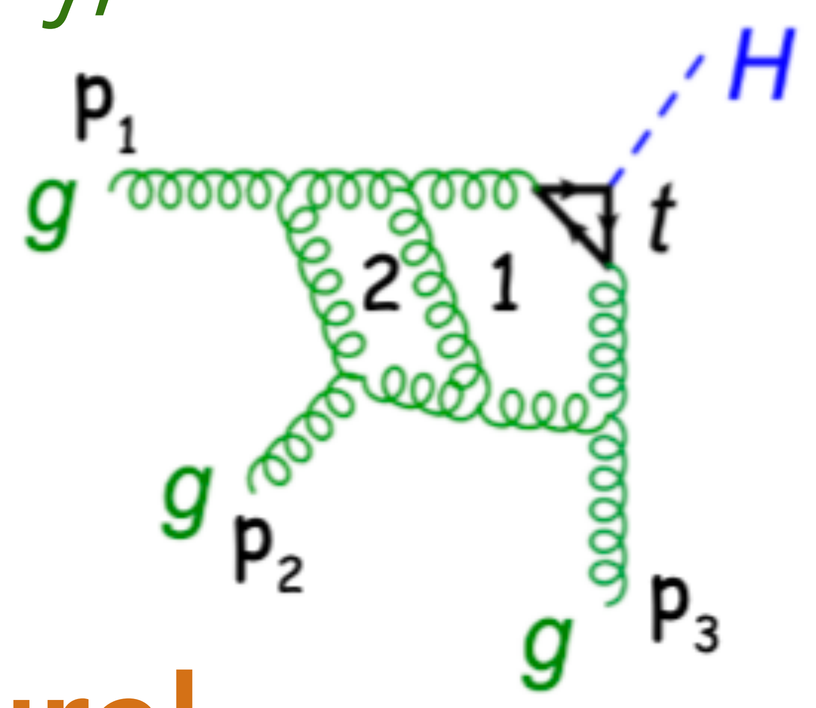
- Of the 6^4 =1296 possible terms, most are 0. **Sparse, lots of structure!**

The Bootstrap

In this bootstrap approach the L-loop amplitude can be expressed as a sum of **terms** with an **integer coefficient** and a **word** composed of 2L **letters**

- The 6 letters $\{a, b, c, d, e, f\}$ encode the kinematics of the collision

For example, in a particular theory called $\mathcal{N}$=4 Super Yang-Mills theory, the answer at 2-loops for a particular interaction is:



$$\mathcal{S}[F_3^{(2)}] = +8bddd +8ceee +8afff +8bfff +8cddd +8aeee$$
$$+16bbbd+16ccce+16bbbf+16aaaf+16cccd+16aaae$$

- Of the 6^4 =1296 possible terms, most are 0. **Sparse, lots of structure!**

The solution space is growing exponentially — **hard to find the answer!**

loop order L	1	2	3	4	5	6	7	8
terms in $S[F_3^{(L)}]$	6	12	636	11,208	263,880	4.9×10^6	9.3×10^7	1.67×10^9

A few SBI Papers in MLST

MACHINE LEARNING
Science and Technology

PAPER • OPEN ACCESS
Stochastic black-box optimization using multi-fidelity score function estimator
Atul Agrawal*, Kislaya Ravi, Phaedon-Stelios Koutsourelakis and Hans-Joachim Bungartz
Published 31 January 2025 • © 2025 The Author(s). Published by IOP Publishing Ltd
Machine Learning: Science and Technology, Volume 6, Number 1
Focus on ML and the Physical Sciences
Citation Atul Agrawal *et al* 2025 *Mach. Learn.: Sci. Technol.* **6** 015024
DOI 10.1088/2632-2153/ad8e2b

OPEN ACCESS
Comparing AI versus optimization workflows for simulation-based inference of spatial-stochastic systems
Michael Alexander Ramirez Sierra* and Thomas R Sokolowski
Published 14 February 2025 • © 2025 The Author(s). Published by IOP Publishing Ltd
Machine Learning: Science and Technology, Volume 6, Number 1
Citation Michael Alexander Ramirez Sierra and Thomas R Sokolowski 2025 *Mach. Learn.: Sci. Technol.* **6** 010502
DOI 10.1088/2632-2153/ada0a3

PAPER • OPEN ACCESS
Simulation-based inference with approximately correct parameters via maximum entropy
Rainier Barrett, Mehrad Ansari, Gourab Ghoshal and Andrew D White
Published 27 April 2022 • © 2022 The Author(s). Published by IOP Publishing Ltd
Machine Learning: Science and Technology, Volume 3, Number 2
Citation Rainier Barrett *et al* 2022 *Mach. Learn.: Sci. Technol.* **3** 025006
DOI 10.1088/2632-2153/ac6286

LETTER • OPEN ACCESS
DIGS: deep inference of galaxy spectra with neural posterior estimation
Gourav Khullar, Brian Nord, Aleksandra Ćiprijanović, Jason Poh and Fei Xu
Published 28 December 2022 • © 2022 The Author(s). Published by IOP Publishing Ltd
Machine Learning: Science and Technology, Volume 3, Number 4
Citation Gourav Khullar *et al* 2022 *Mach. Learn.: Sci. Technol.* **3** 04LT04
DOI 10.1088/2632-2153/ac98f4

PAPER • OPEN ACCESS
Evidence Networks: simple losses for fast, amortized, neural Bayesian model comparison
Niall Jeffrey and Benjamin D Wandelt
Published 17 January 2024 • © 2024 The Author(s). Published by IOP Publishing Ltd
Machine Learning: Science and Technology, Volume 5, Number 1
Citation Niall Jeffrey and Benjamin D Wandelt 2024 *Mach. Learn.: Sci. Technol.* **5** 015008
DOI 10.1088/2632-2153/ad1a4d

PAPER • OPEN ACCESS
Simulation-based inference on virtual brain models of disorders
Meysam Hashemi, Abolfazl Ziaeemehr, Marmaduke M Woodman, Jan Fousek, Spase Petkoski and Viktor K Jirsa
Published 19 July 2024 • © 2024 The Author(s). Published by IOP Publishing Ltd
Machine Learning: Science and Technology, Volume 5, Number 3
Focus on Explainable Machine Learning in Sciences
Citation Meysam Hashemi *et al* 2024 *Mach. Learn.: Sci. Technol.* **5** 035019
DOI 10.1088/2632-2153/ad6230

PAPER • OPEN ACCESS
Importance nested sampling with normalising flows
Michael J Williams, John Veitch and Chris Messenger
Published 25 July 2023 • © 2023 The Author(s). Published by IOP Publishing Ltd
Machine Learning: Science and Technology, Volume 4, Number 3
Citation Michael J Williams *et al* 2023 *Mach. Learn.: Sci. Technol.* **4** 035011
DOI 10.1088/2632-2153/acd5aa

PAPER • OPEN ACCESS
Multi-fidelity Gaussian process surrogate modeling for regression problems in physics
Kislaya Ravi*, Vladyslav Fediukov*, Felix Dietrich, Tobias Neckel, Fabian Buse, Michael Bergmann and Hans-Joachim Bungartz
Published 15 October 2024 • © 2024 The Author(s). Published by IOP Publishing Ltd
Machine Learning: Science and Technology, Volume 5, Number 4
Focus on ML and the Physical Sciences
Citation Kislaya Ravi *et al* 2024 *Mach. Learn.: Sci. Technol.* **5** 045015
DOI 10.1088/2632-2153/ad7ad5

PAPER • OPEN ACCESS
Efficient Bayesian inference using physics-informed invertible neural networks for inverse problems
Xiaofei Guan, Xintong Wang, Hao Wu, Zihao Yang and Peng Yu
Published 23 July 2024 • © 2024 The Author(s). Published by IOP Publishing Ltd
Machine Learning: Science and Technology, Volume 5, Number 3
Focus on Generative AI in Science
Citation Xiaofei Guan *et al* 2024 *Mach. Learn.: Sci. Technol.* **5** 035026
DOI 10.1088/2632-2153/ad5f74

PAPER • OPEN ACCESS
`DiffLense`: a conditional diffusion model for super-resolution of gravitational lensing data
Pranath Reddy*, Michael W Toomey, Hanna Parul and Sergei Gleyzer
Published 19 September 2024 • © 2024 The Author(s). Published by IOP Publishing Ltd
Machine Learning: Science and Technology, Volume 5, Number 3
Focus on ML and the Physical Sciences
Citation Pranath Reddy *et al* 2024 *Mach. Learn.: Sci. Technol.* **5** 035076
DOI 10.1088/2632-2153/ad76f8

Unfolding

EUROPEAN ORGANISATION FOR NUCLEAR RESEARCH (CERN)

ATLAS EXPERIMENT

Phys.Rev.Lett. 133 (2024) 26, 261803
DOI: 10.1103/PhysRevLett.133.261803

CERN-EP-2024-132
February 10, 2025

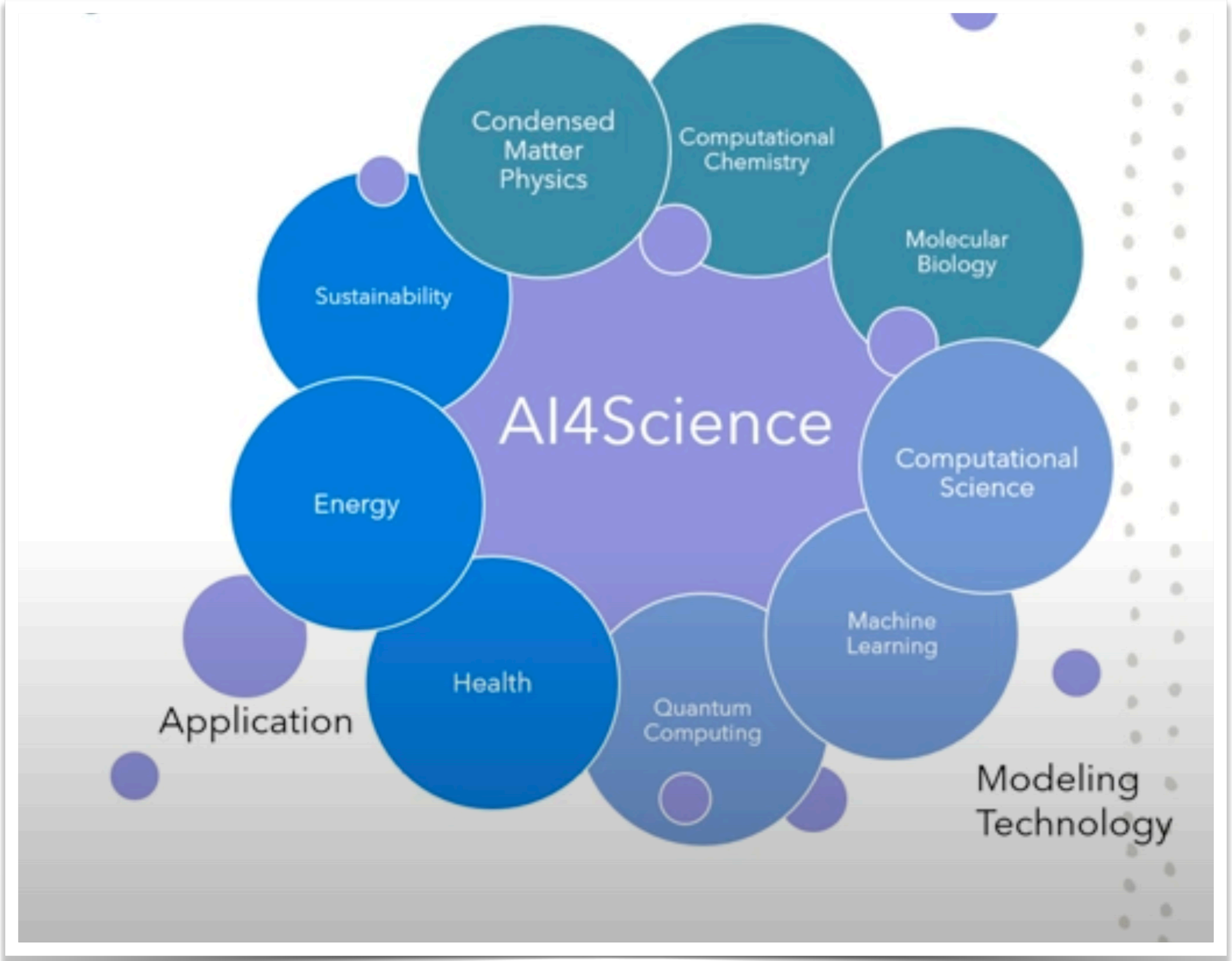
A simultaneous unbinned differential cross section measurement of twenty-four Z+jets kinematic observables with the ATLAS detector

The ATLAS Collaboration

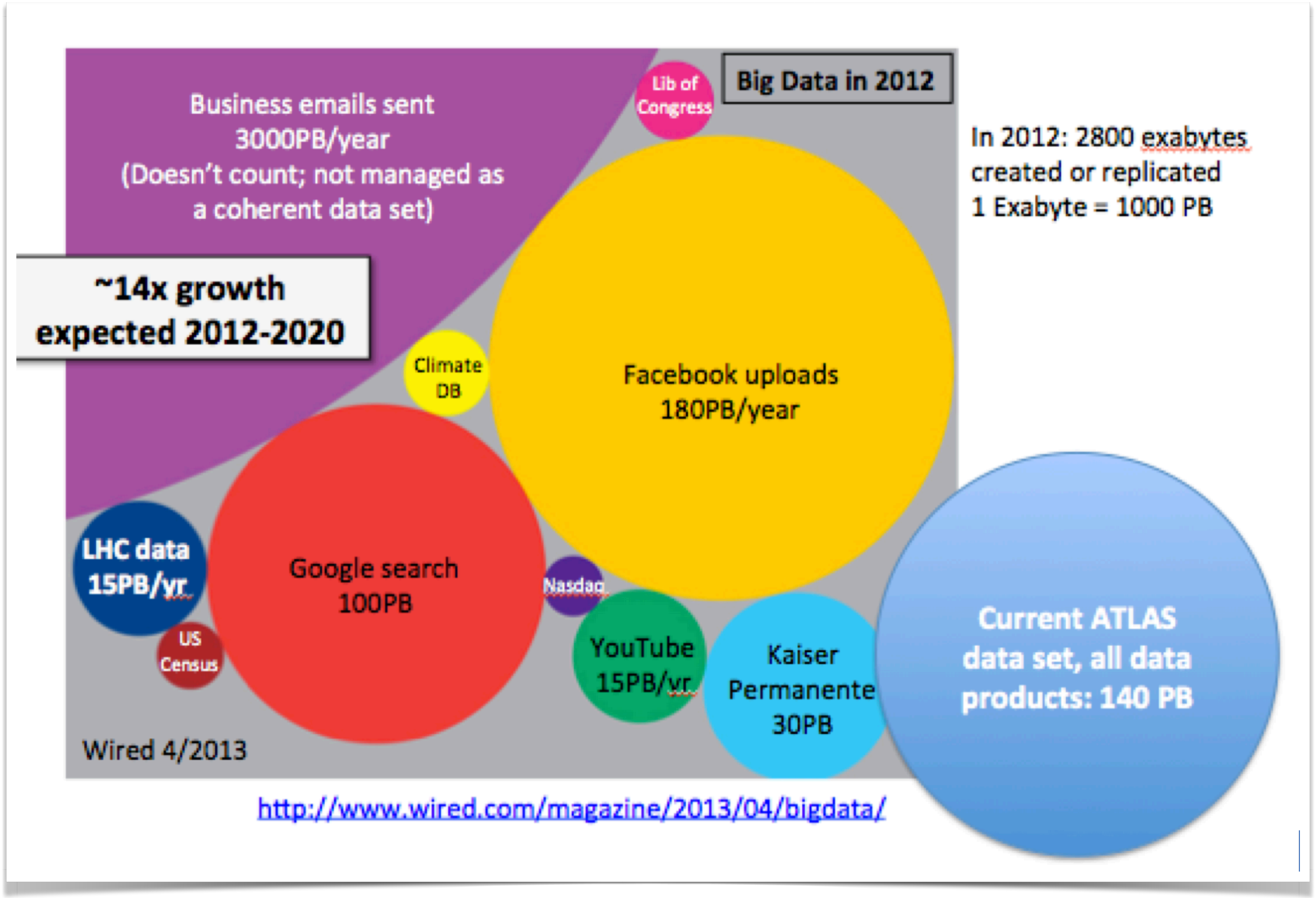
Max Welling's talk @ EuCAIFCon2024

Link to talk

"Where is fundamental physics in AI4Science?" — Lukas Heinrich



Compare with HEP's association with **Big Data**



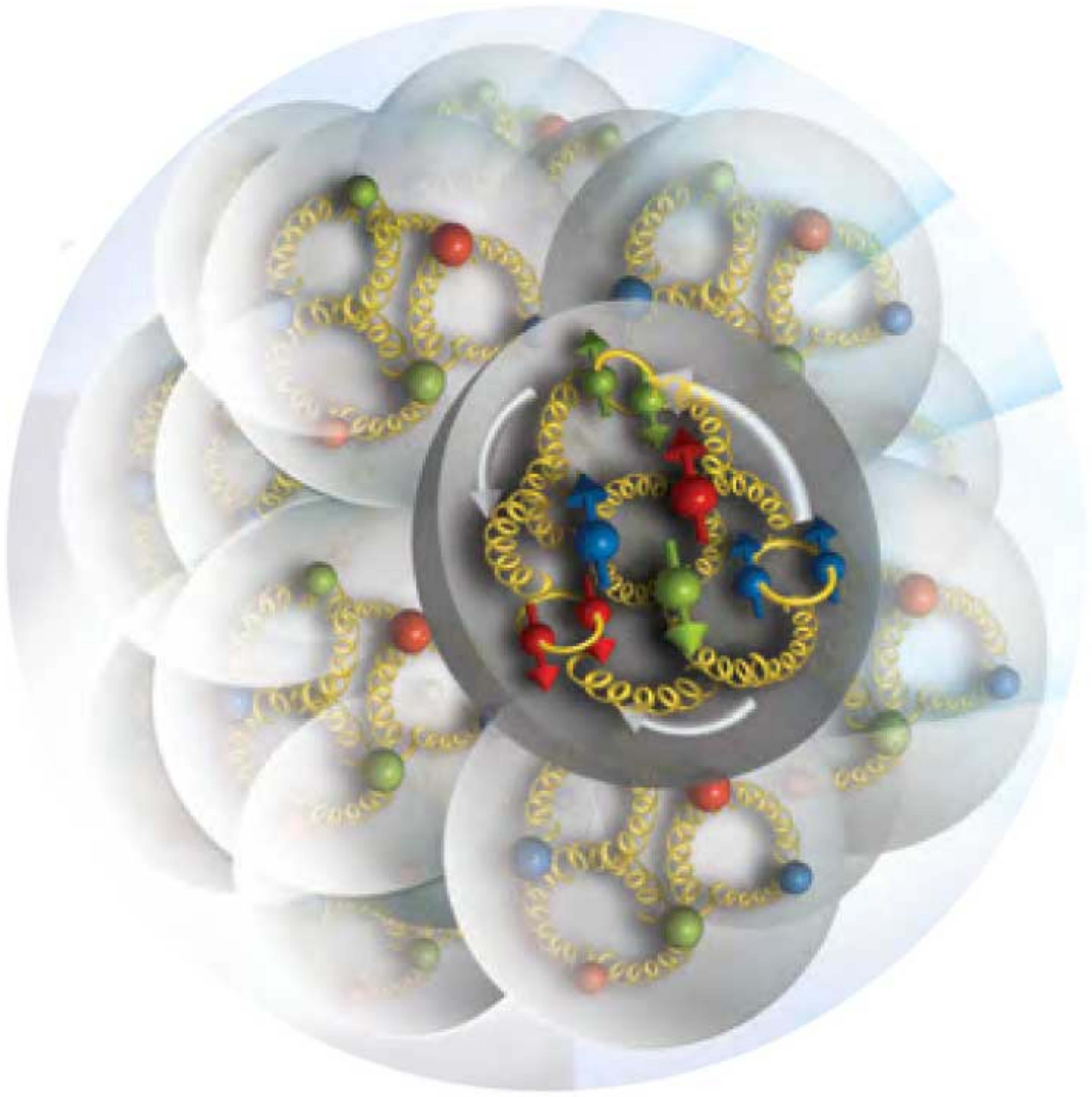
The strong force: Quantum Chromodynamics (QCD)

The strong nuclear force is one of the four fundamental forces.

It is described by Quantum Chromodynamics (QCD)

QCD describes how quarks and gluons interact

Emergent phenomena:
Quarks and gluons form protons, neutrons, etc.

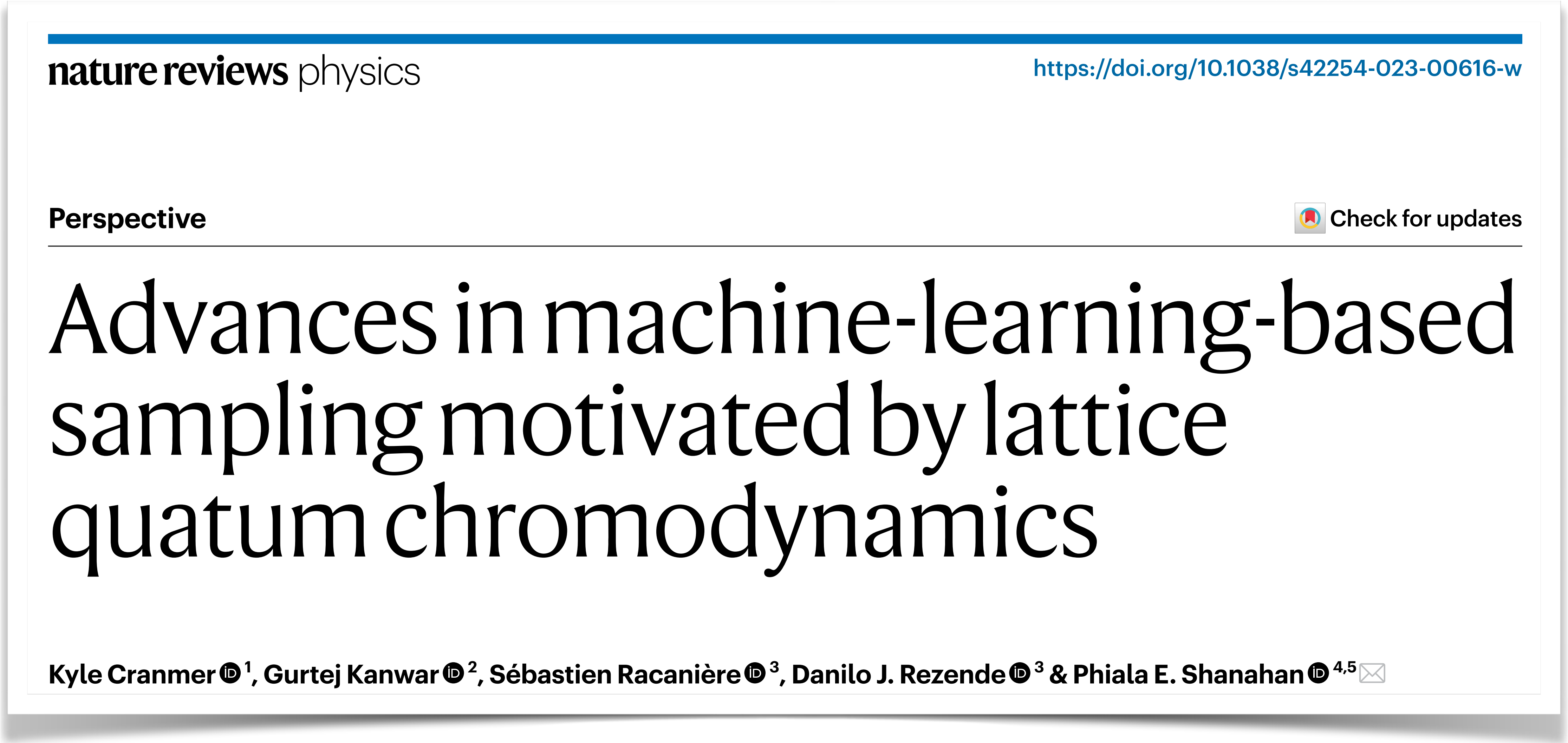


AI-Enhanced Sampling

Basic idea:

- use **generative AI** model (normalizing flows) to approximate the target Boltzmann distribution.
- Train using reverse KL[q||p] (not samples from the target)
- Sample from the flow instead of traditional Hamiltonian MC

Learned model won't be perfect, but you can **correct** via importance sampling or MCMC procedure


nature reviews physics

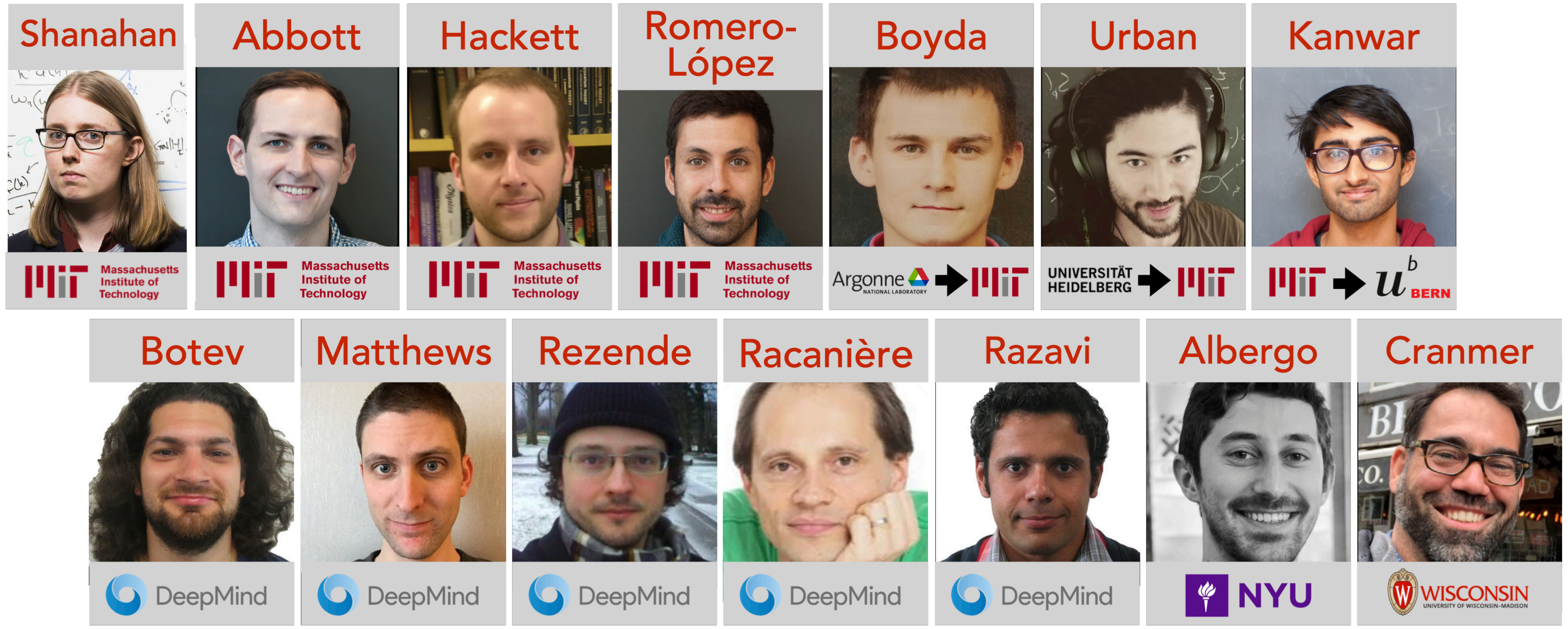
https://doi.org/10.1038/s42254-023-00616-w

Perspective

Check for updates

Advances in machine-learning-based sampling motivated by lattice quatum chromodynamics

Kyle Cranmer [1], Gurtej Kanwar [2], Sébastien Racanière [3], Danilo J. Rezende [3] & Phiala E. Shanahan [4,5]

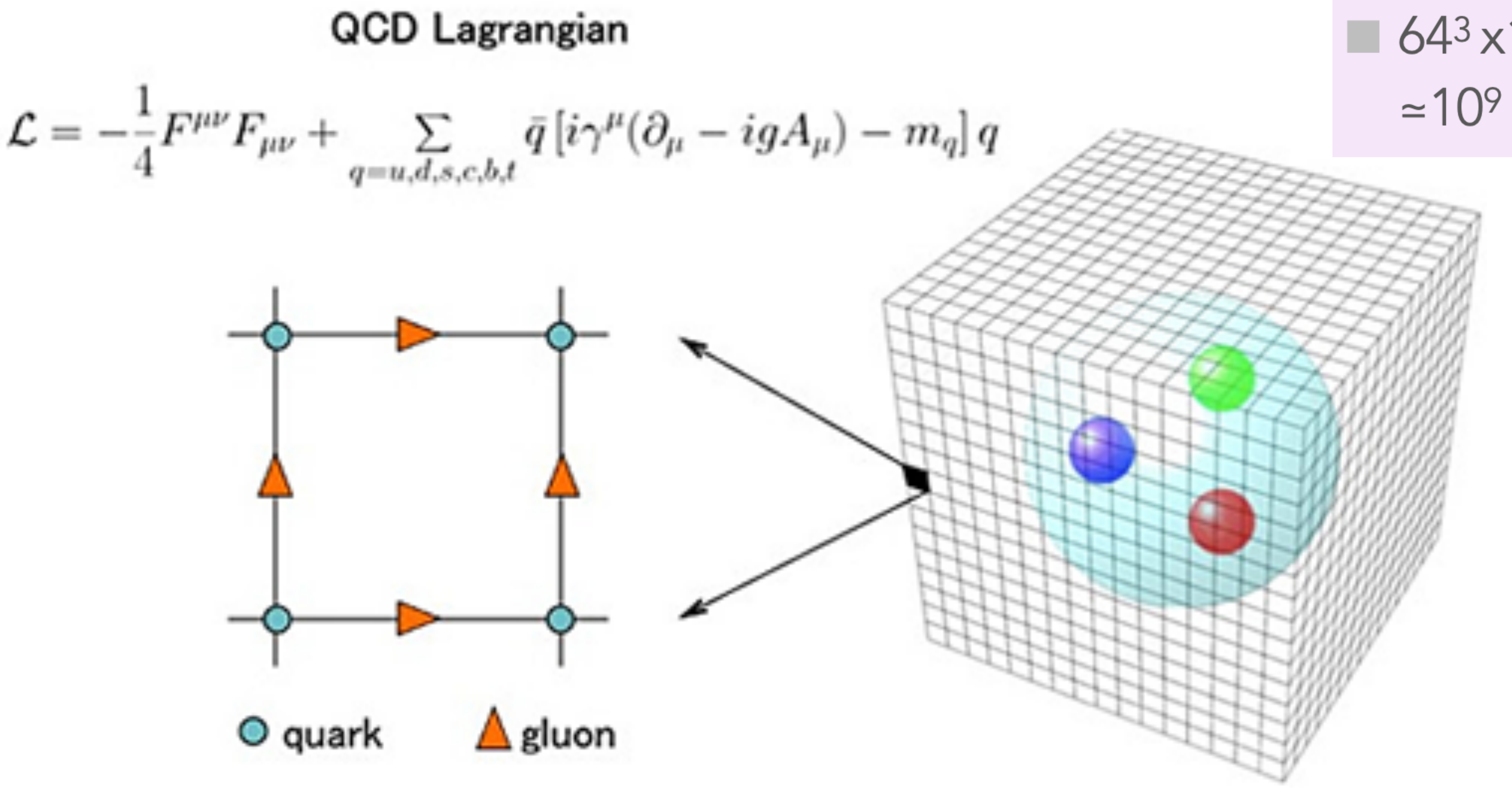


See also: Albergo, Kanwar, Shanahan, PRD (2019) arXiv:1904.12072

Lattice Field Theory

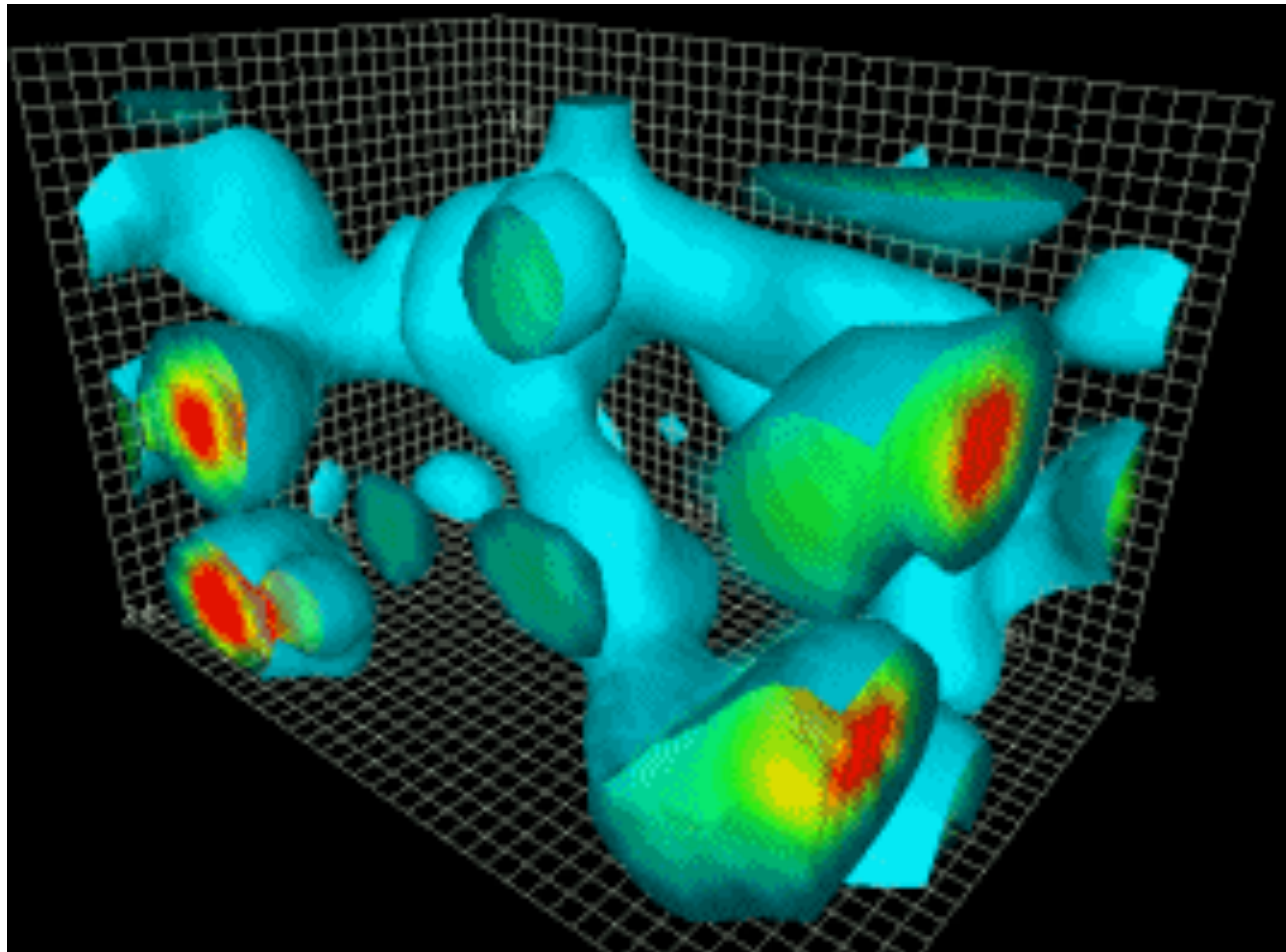
Lattice field theory is a computational approach to studying interacting field theory on a discretized space-time lattice.

Each link on the lattice has data corresponding to the symmetry group of the theory. For the strong force (QCD) each link has a 3x3 unitary matrix.



- $64^3 \times 128 \times 4 \times 9 \times 2$
  $\simeq 10^9$ numbers

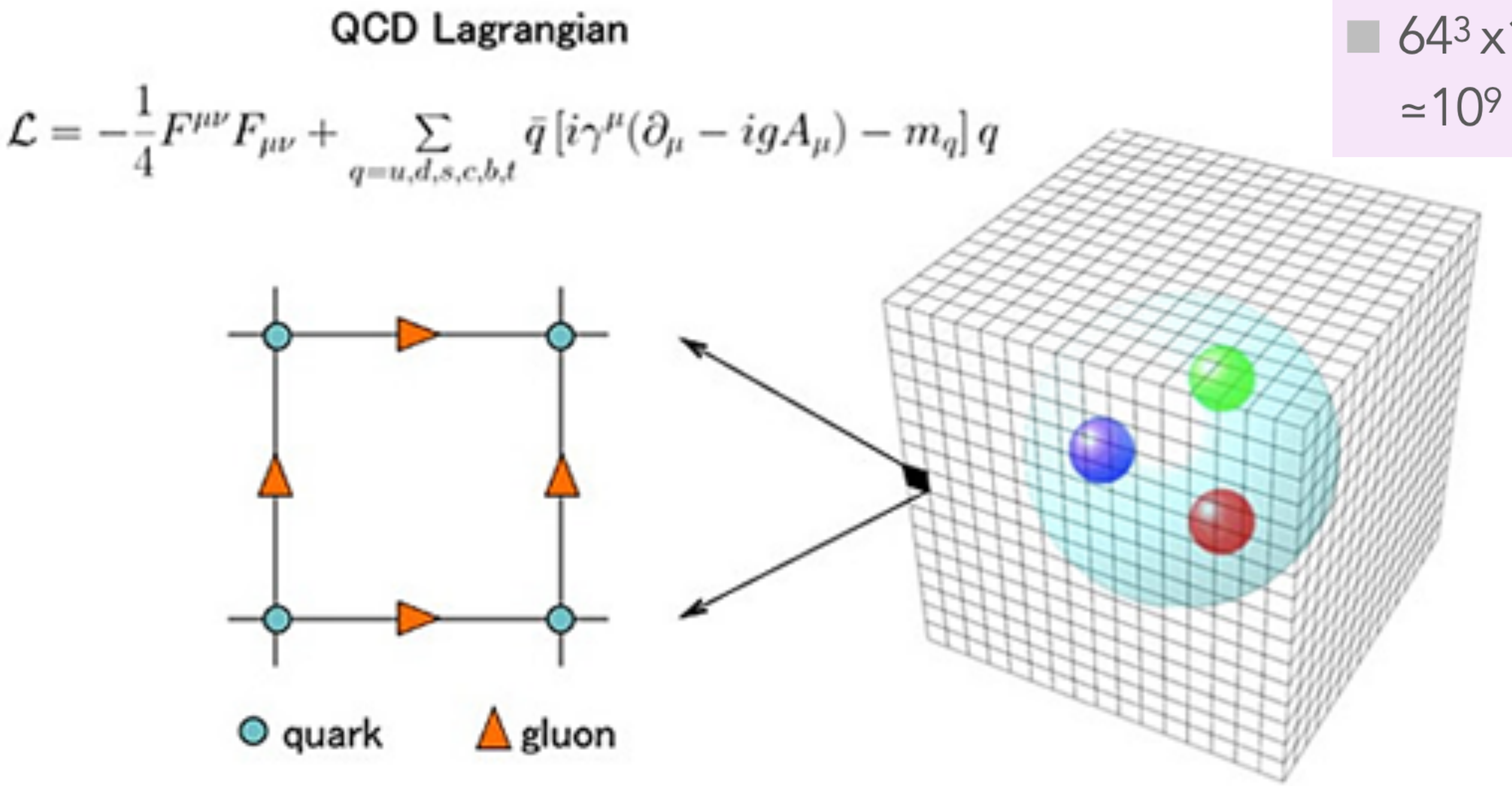
This animation is a single configuration of the lattice. Think of a 4-d image playing like a movie.



Lattice Field Theory

Lattice field theory is a computational approach to studying interacting field theory on a discretized space-time lattice.

Each link on the lattice has data corresponding to the symmetry group of the theory. For the strong force (QCD) each link has a 3x3 unitary matrix.



- $64^3 \times 128 \times 4 \times 9 \times 2$ $\simeq 10^9$ numbers

This animation is a single configuration of the lattice. Think of a 4-d image playing like a movie.

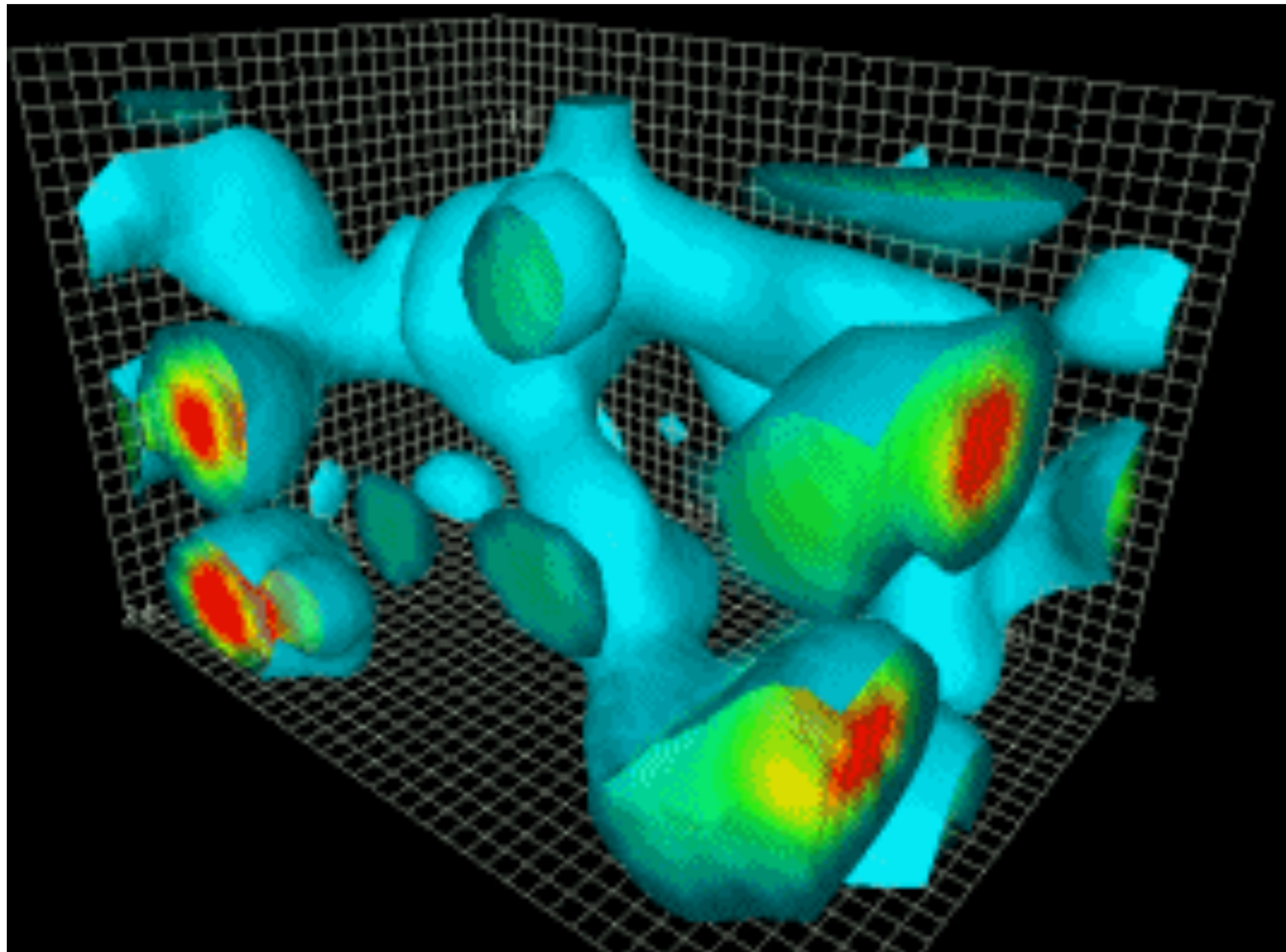


Image vs. Lattice Quantum Fields

Image generation

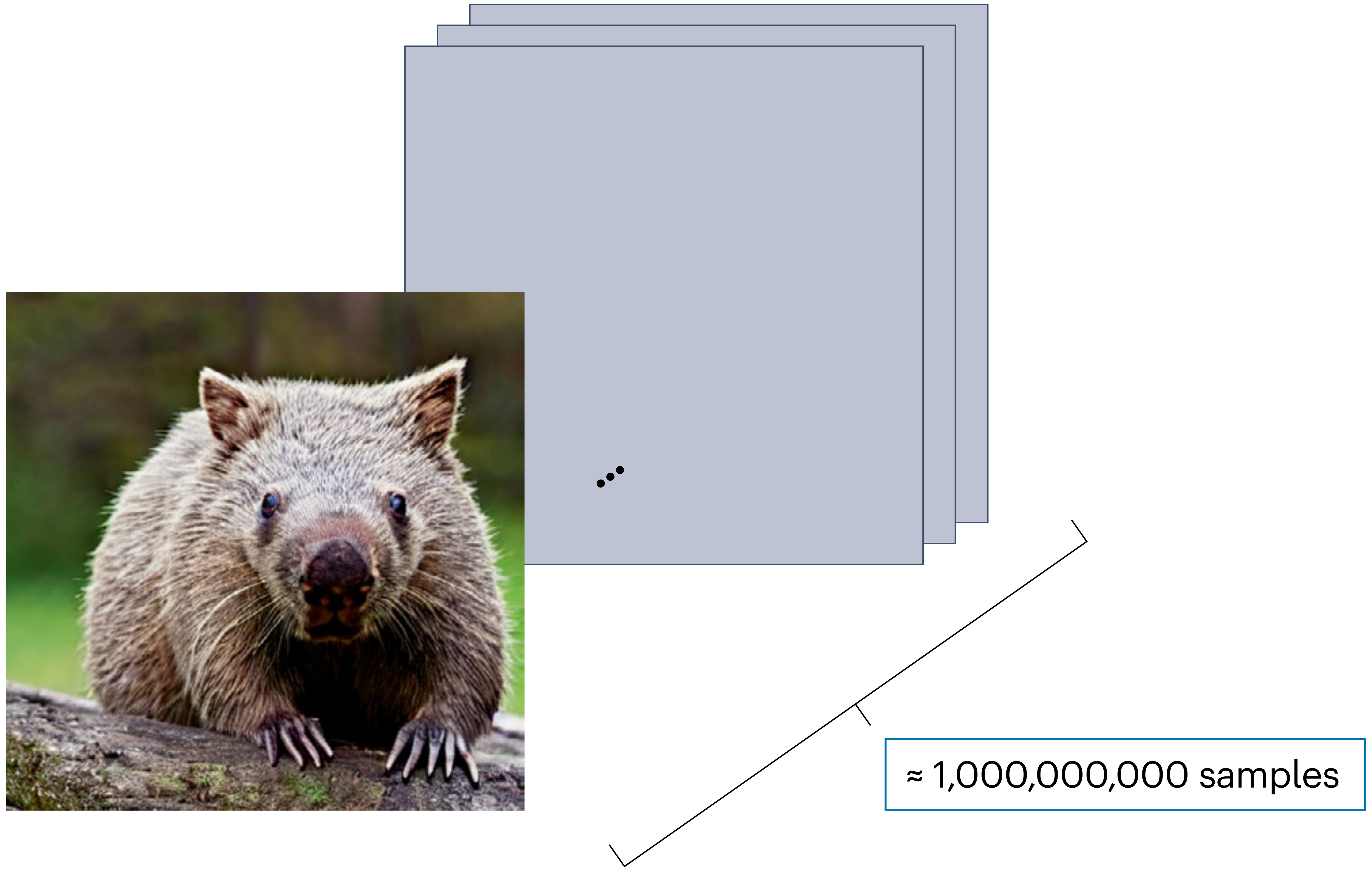


Image geometry	**512 × 512**
RGB pixel variables	**× 3**
	≈ 1,000,000 dof

Target
Subjective high quality per sample

Symmetries
Few approximate symmetries
(for example, reflection, small translations)

Image vs. Lattice Quantum Fields

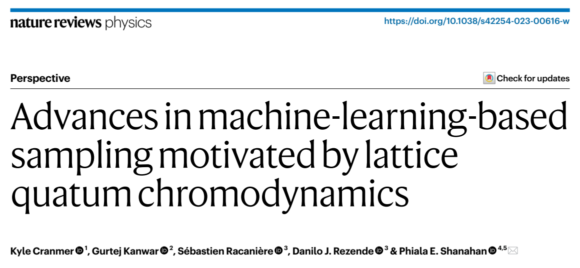

nature reviews physics
Perspective
Advances in machine-learning-based sampling motivated by lattice quatum chromodynamics

Image generation

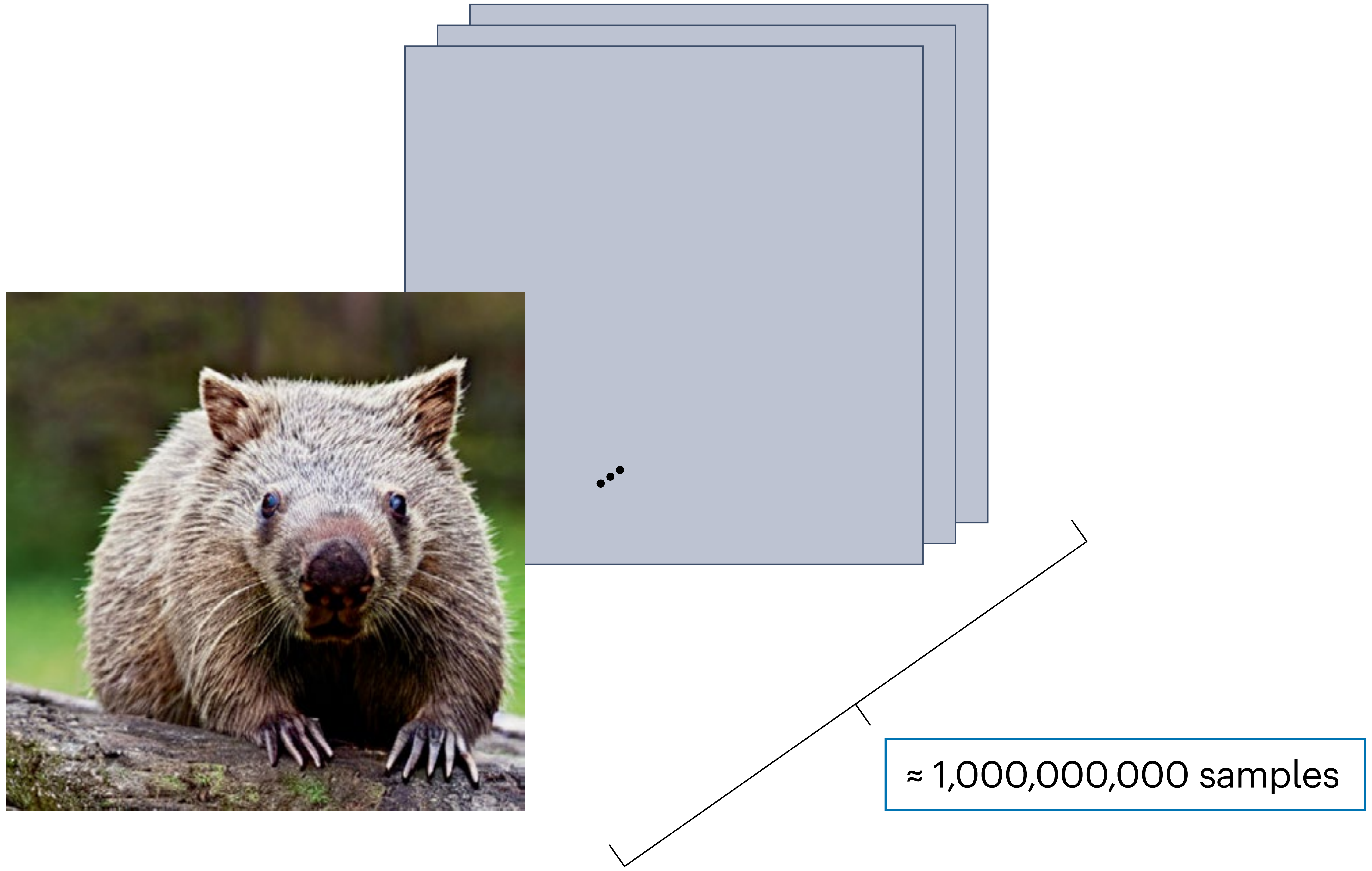
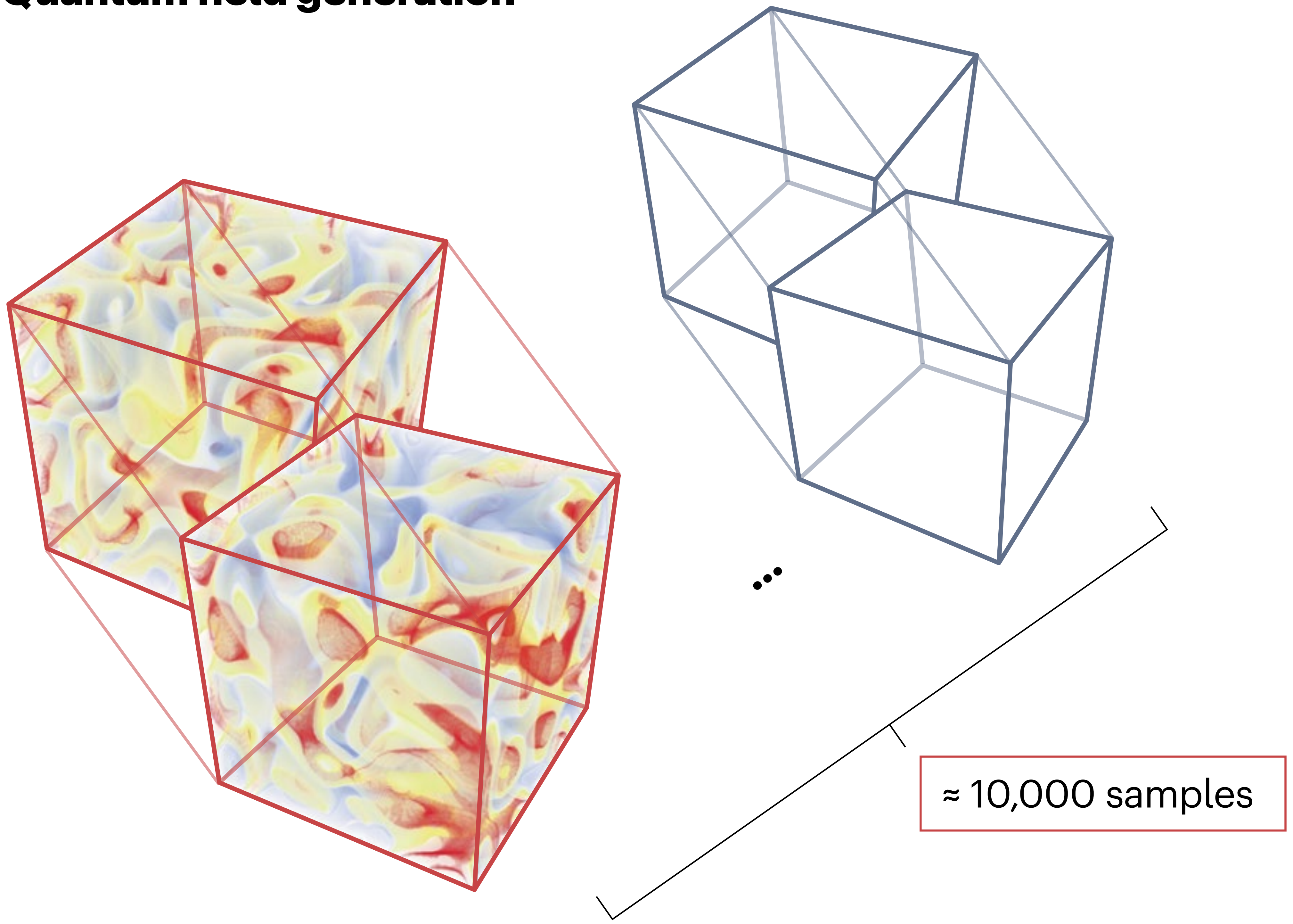


Image geometry	**512 × 512**
RGB pixel variables	**× 3**
	≈ 1,000,000 dof

Target
Subjective high quality per sample

Symmetries
Few approximate symmetries
(for example, reflection, small translations)

Quantum field generation



Lattice geometry	**256 × 256 × 256 × 512**
SU(3) link variables	**× 4 × 8**
	≈ 100,000,000,000 dof

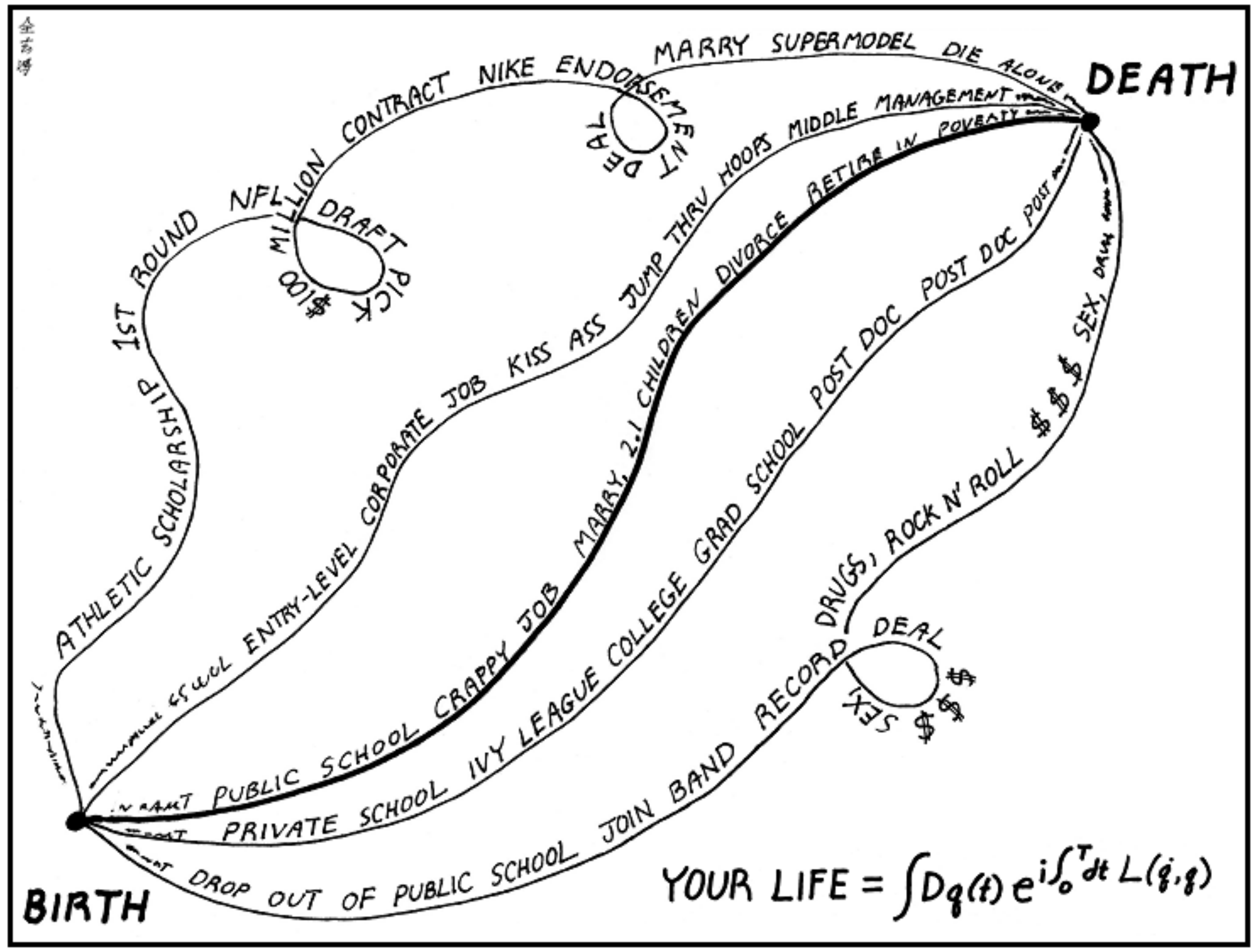
Target
Objective distribution $p(U) = e^{-S(U)}/Z$

Symmetries
High-dimensional exact symmetries
(for example, translations, gauge symmetry)

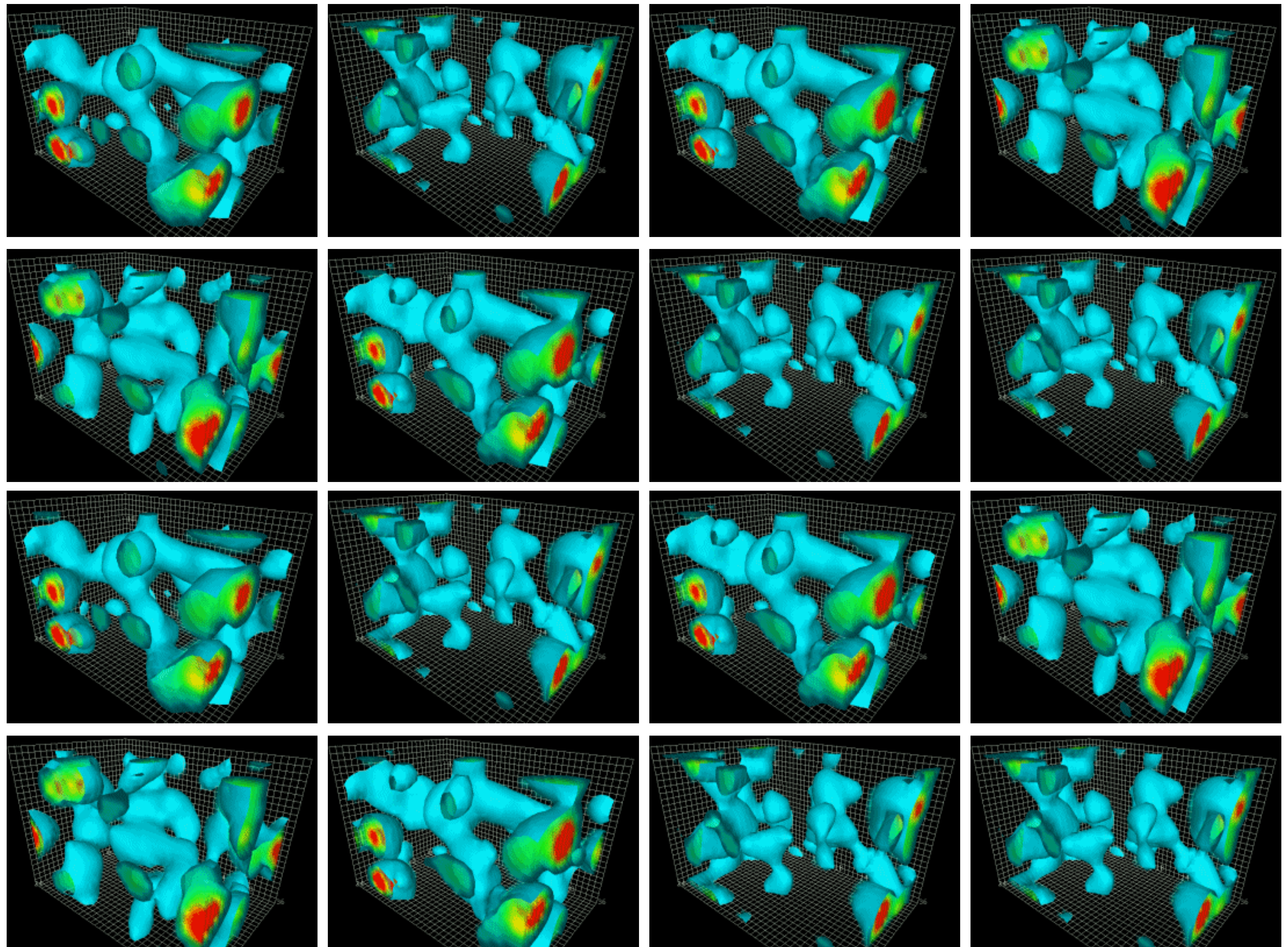
Distribution over configurations

We don't want just a single "image" (lattice configuration), we want to **sample** the high-dimensional distribution of configurations predicted by the theory.

- **Path integral**: each "path" is a sample from distribution of lattice configurations path **~exp(-Action[path])**
- Predictions are **expectations** of quantum operators w.r.t. this distribution.
- **Hamiltonian Monte Carlo was invented for this problem**, but it has limitations.



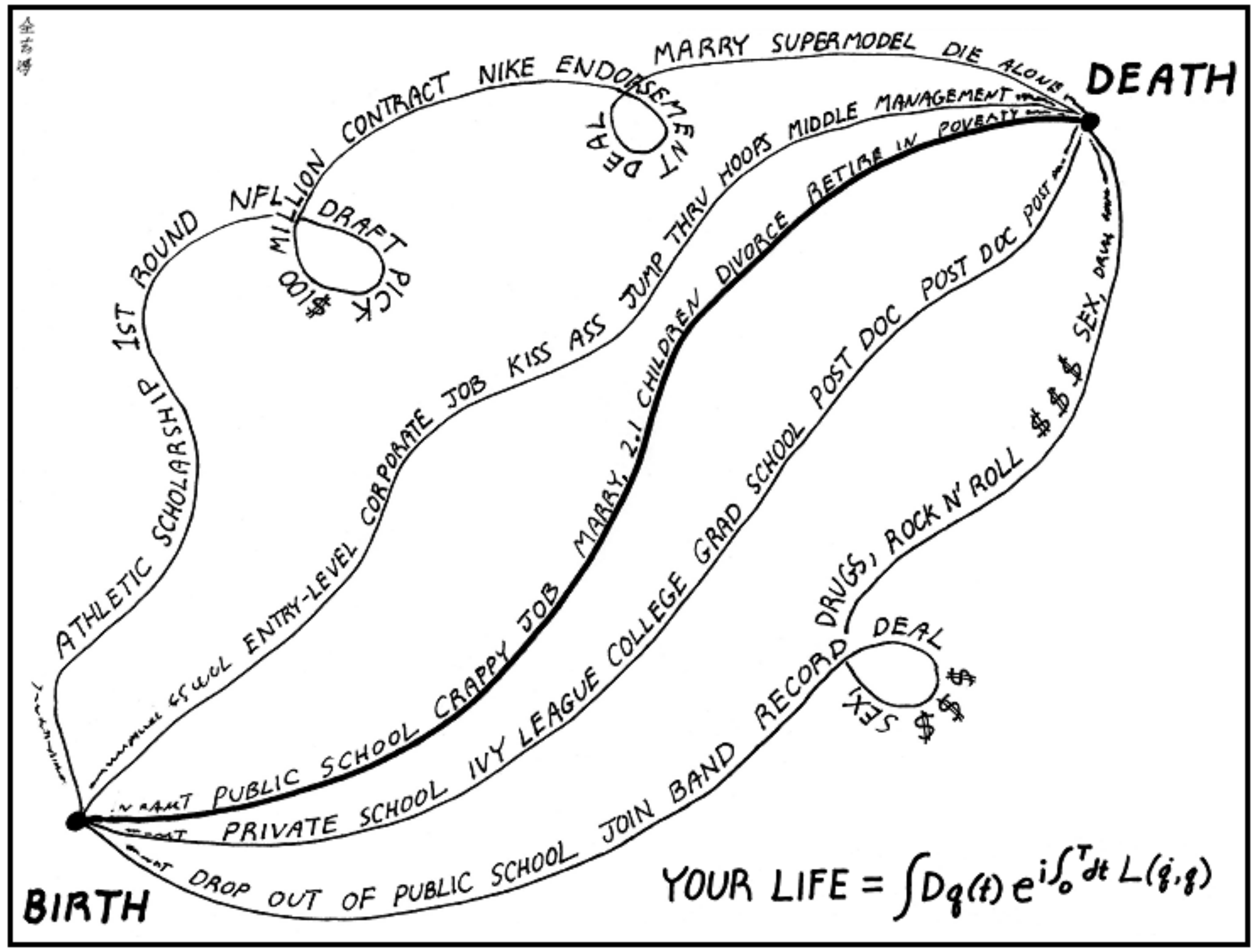
The Path Integral Formulation of Your Life



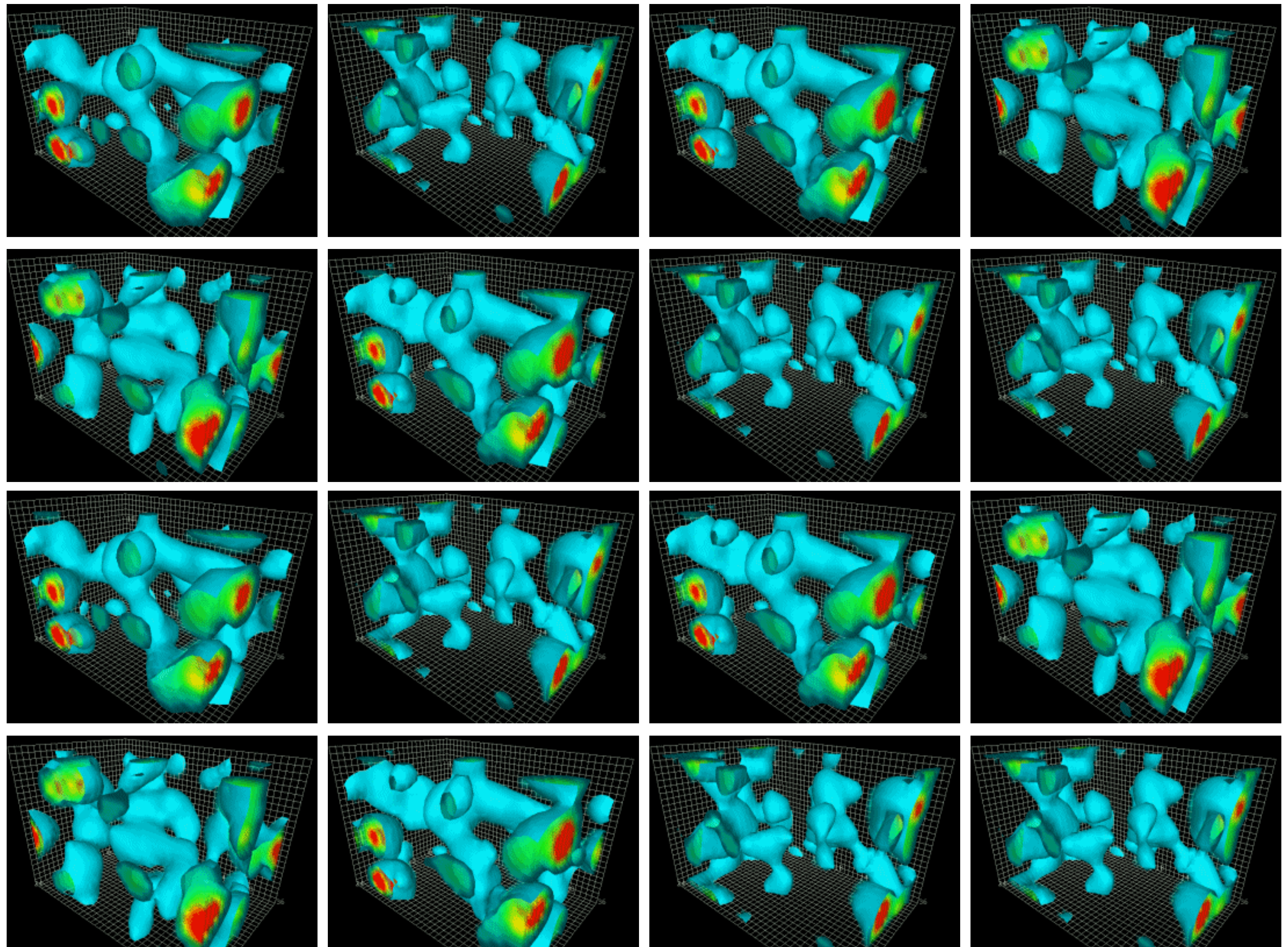
Distribution over configurations

We don't want just a single "image" (lattice configuration), we want to **sample** the high-dimensional distribution of configurations predicted by the theory.

- **Path integral**: each "path" is a sample from distribution of lattice configurations path **~exp(-Action[path])**
- Predictions are **expectations** of quantum operators w.r.t. this distribution.
- **Hamiltonian Monte Carlo was invented for this problem**, but it has limitations.

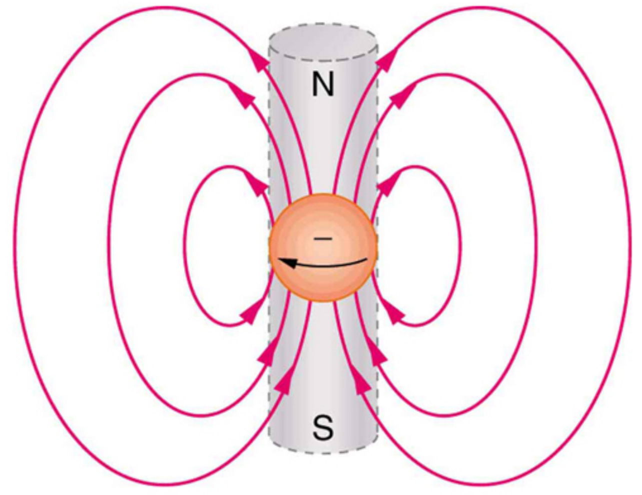


The Path Integral Formulation of Your Life

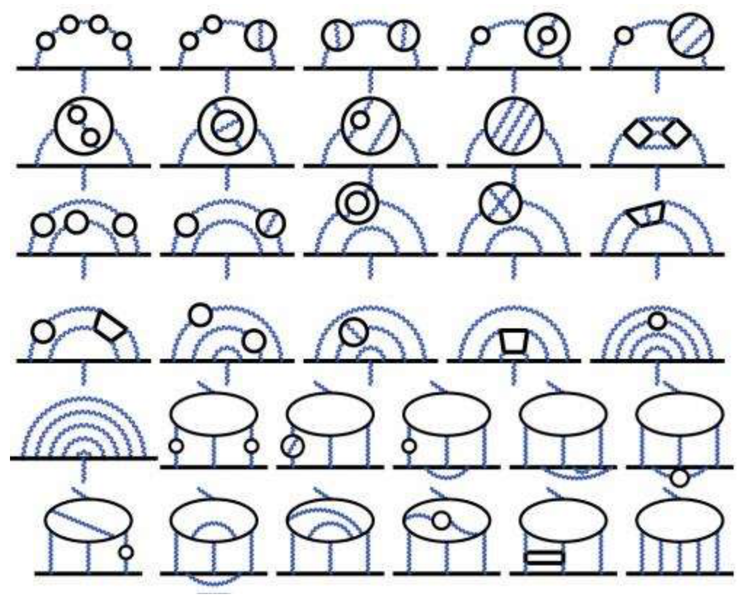


Predictions are taken seriously

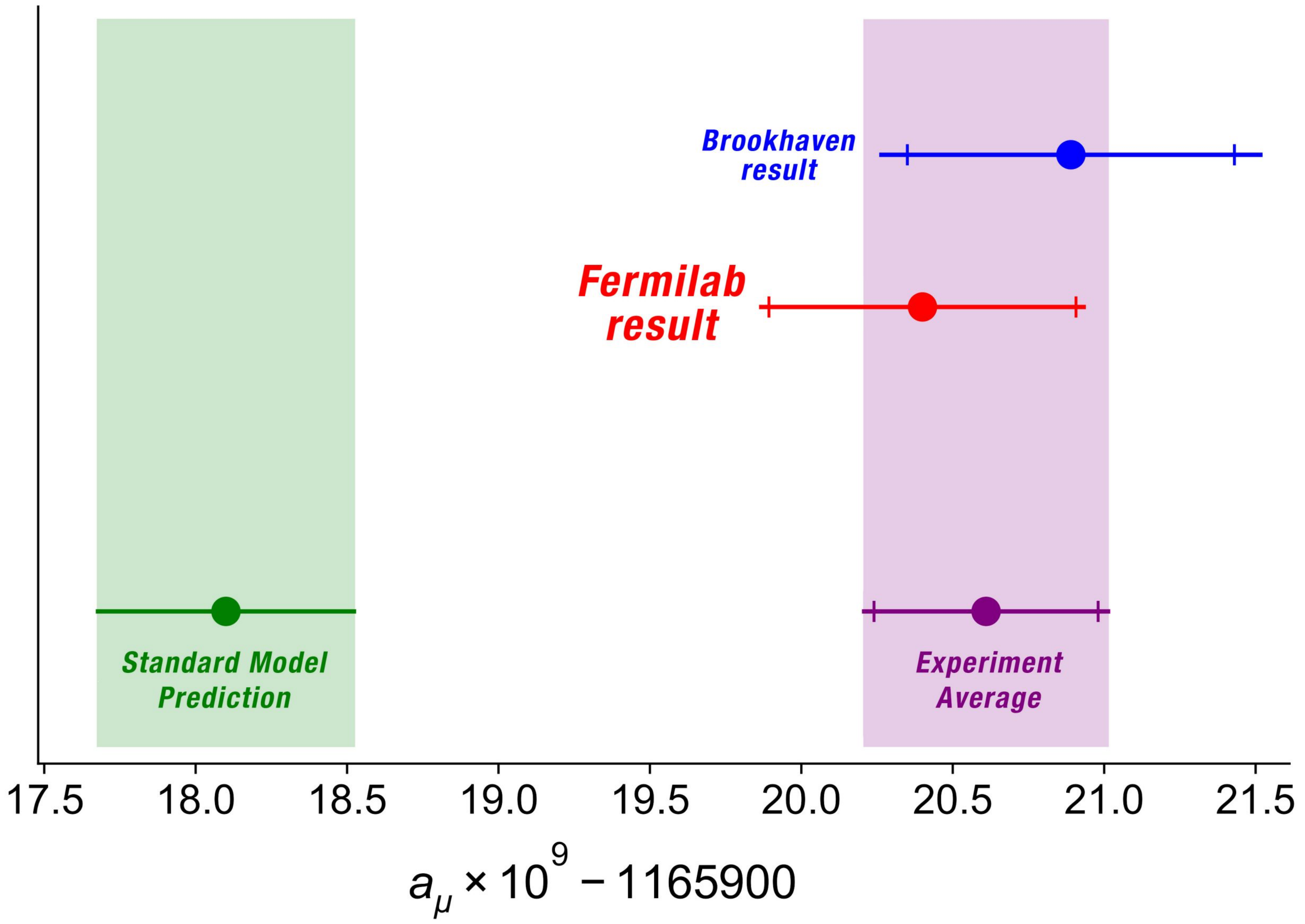
Magnetic moment of the electron:
(torque an electron feels in a magnetic field) $a_e = (g-2)/2$

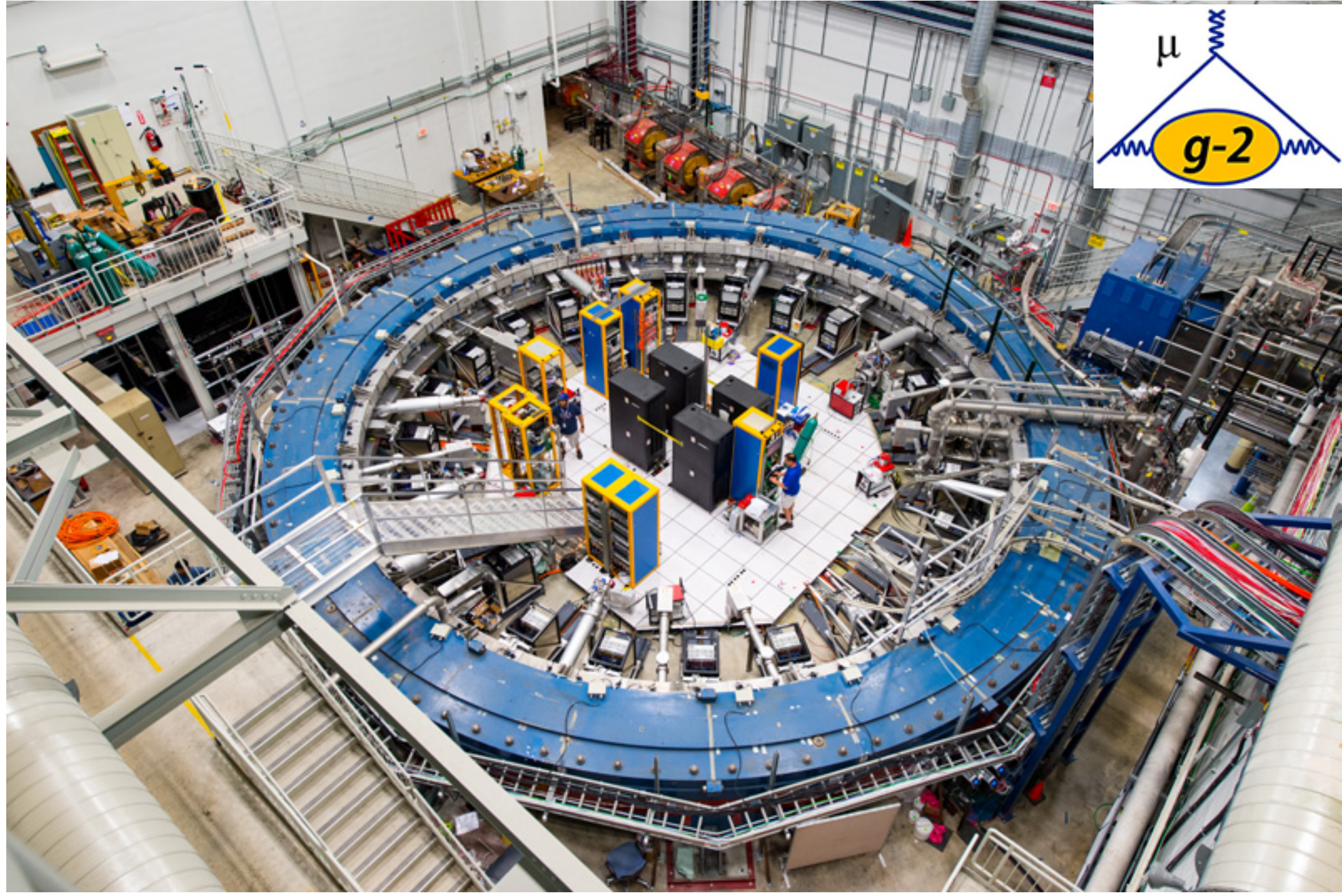


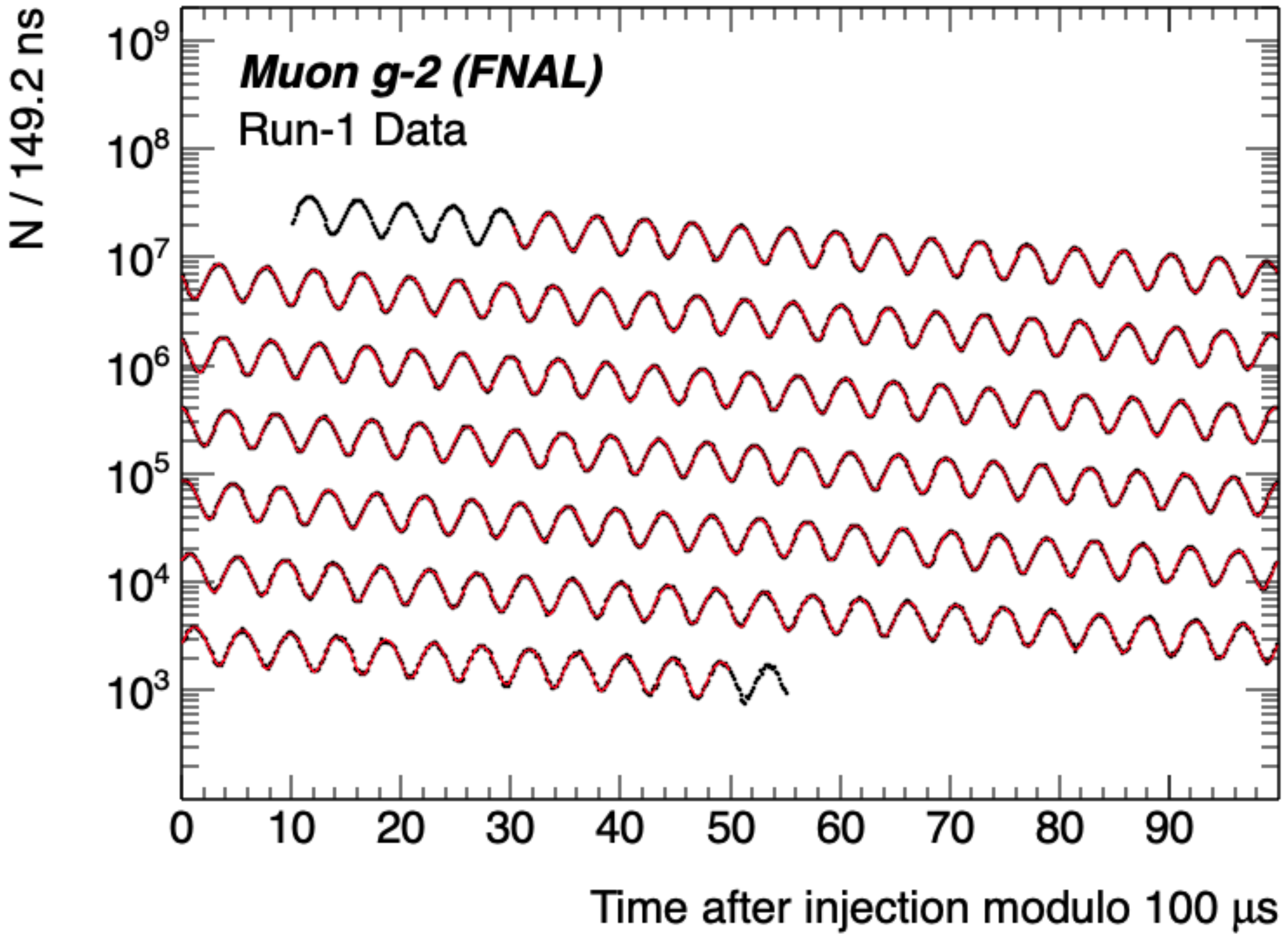
Most accurately verified prediction in the history of physics



Theory	$a_e = 0.001159652181643(764)$
Exp.	$a_e = 0.00115965218073(28)$







Flows for LQCD

Albergo, Kanwar, Shanahan, PRD (2019) arXiv:1904.12072

Basic idea:

- use **normalizing flows** to approximate the target Boltzmann distribution.
- Train using reverse KL[q||p] (not samples from the target)
- Sample from the flow instead of traditional Hamiltonian MC

Learned model won't be perfect, but you can **correct** via importance sampling or MCMC procedure

Flow-based generative models for Markov chain Monte Carlo in lattice field theory

M. S. Albergo,[1,2,3] G. Kanwar,[4] and P. E. Shanahan[4,1]

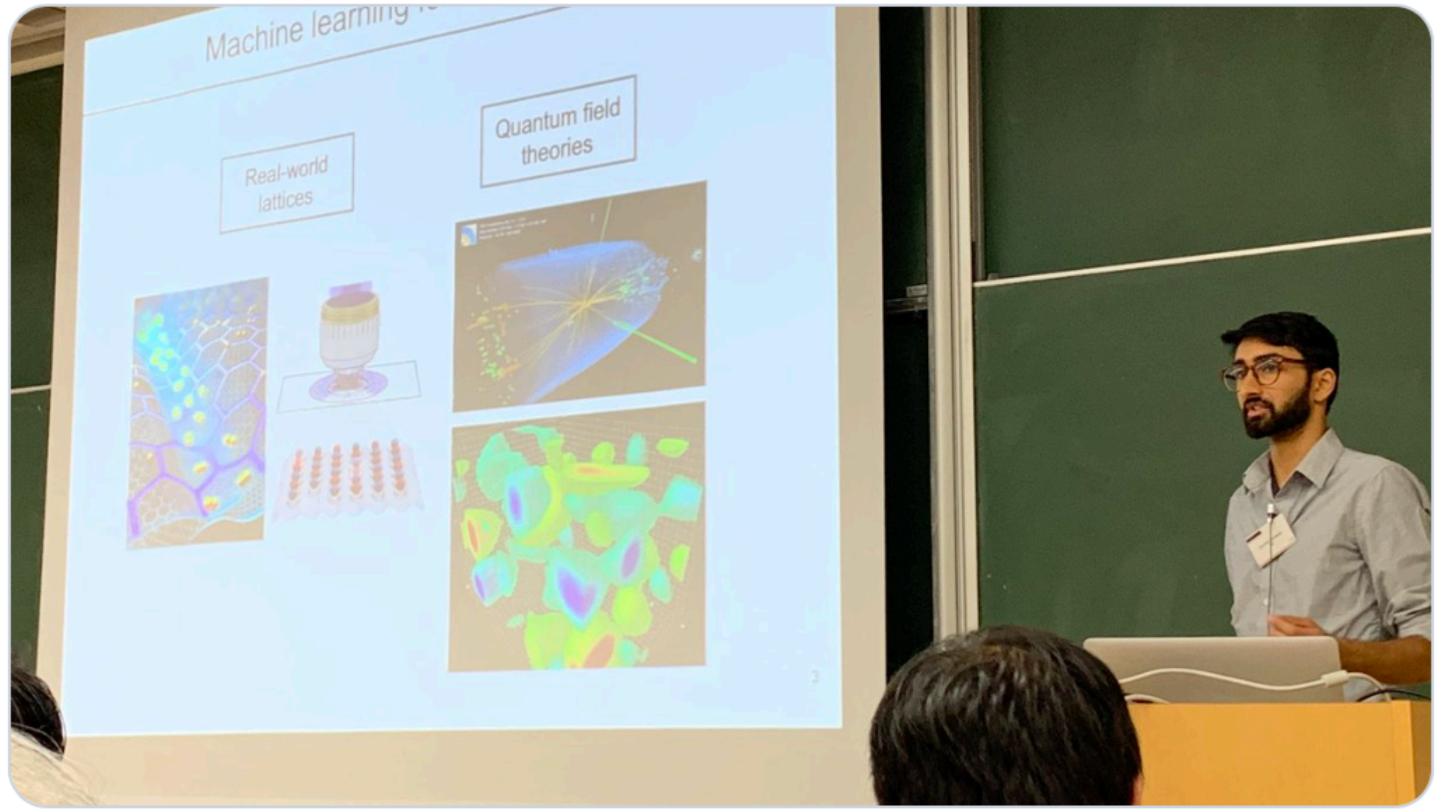
[1] *Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2L 2Y5, Canada*
[2] *Cavendish Laboratories, University of Cambridge, Cambridge CB3 0HE, U.K.*
[3] *University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*
[4] *Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, U.S.A.*

A Markov chain update scheme using a machine-learned *flow-based generative model* is proposed for Monte Carlo sampling in lattice field theories. The generative model may be optimized (trained) to produce samples from a distribution approximating the desired Boltzmann distribution determined by the lattice action of the theory being studied. Training the model systematically improves autocorrelation times in the Markov chain, even in regions of parameter space where standard Markov chain Monte Carlo algorithms exhibit critical slowing down in producing decorrelated updates. Moreover, the model may be trained without existing samples from the desired distribution. The algorithm is compared with HMC and local Metropolis sampling for ϕ^4 theory in two dimensions.

Hire Michael ↑

Enrico Rinaldi @enricesena · Nov 1

Yesterday **Gurtej Kanwar** told us about machine learning for lattice field theories and exciting progress in Generative Models for gauge theories (collaboration with @DeepMindAI) at #DLAP2019 Today is the last day of this great conference!



5 15