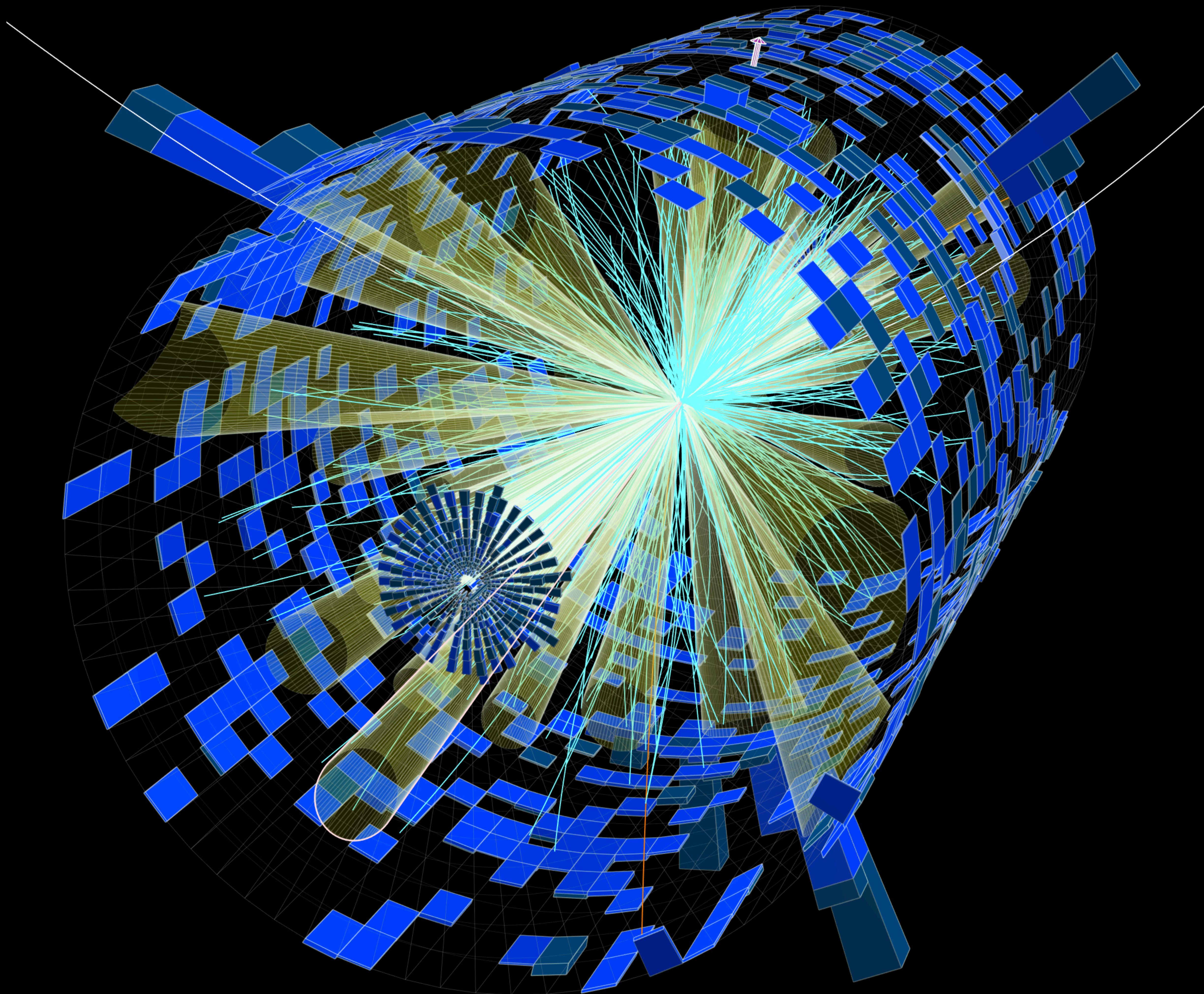
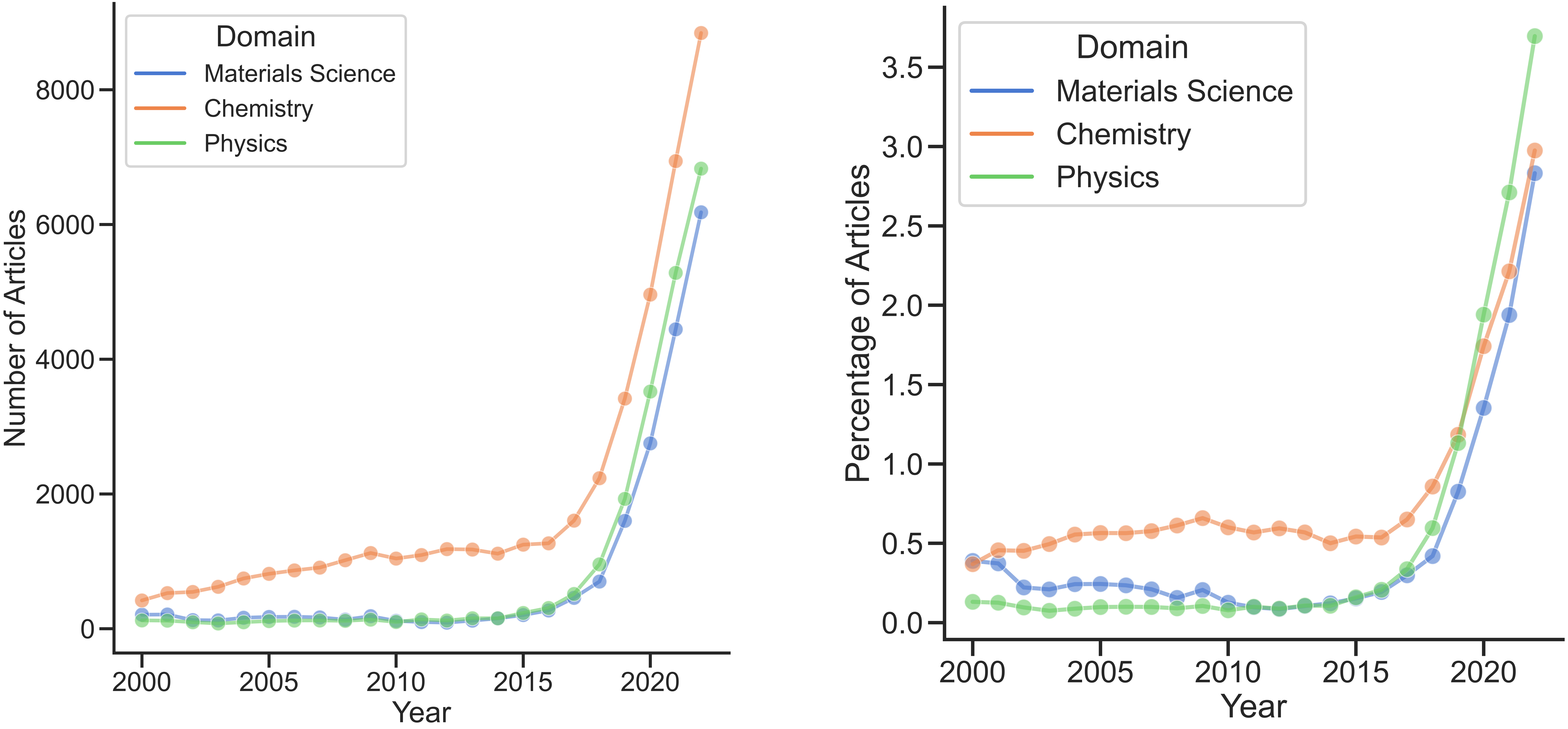
AI for Fundamental Physics



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

ML Publications in Science

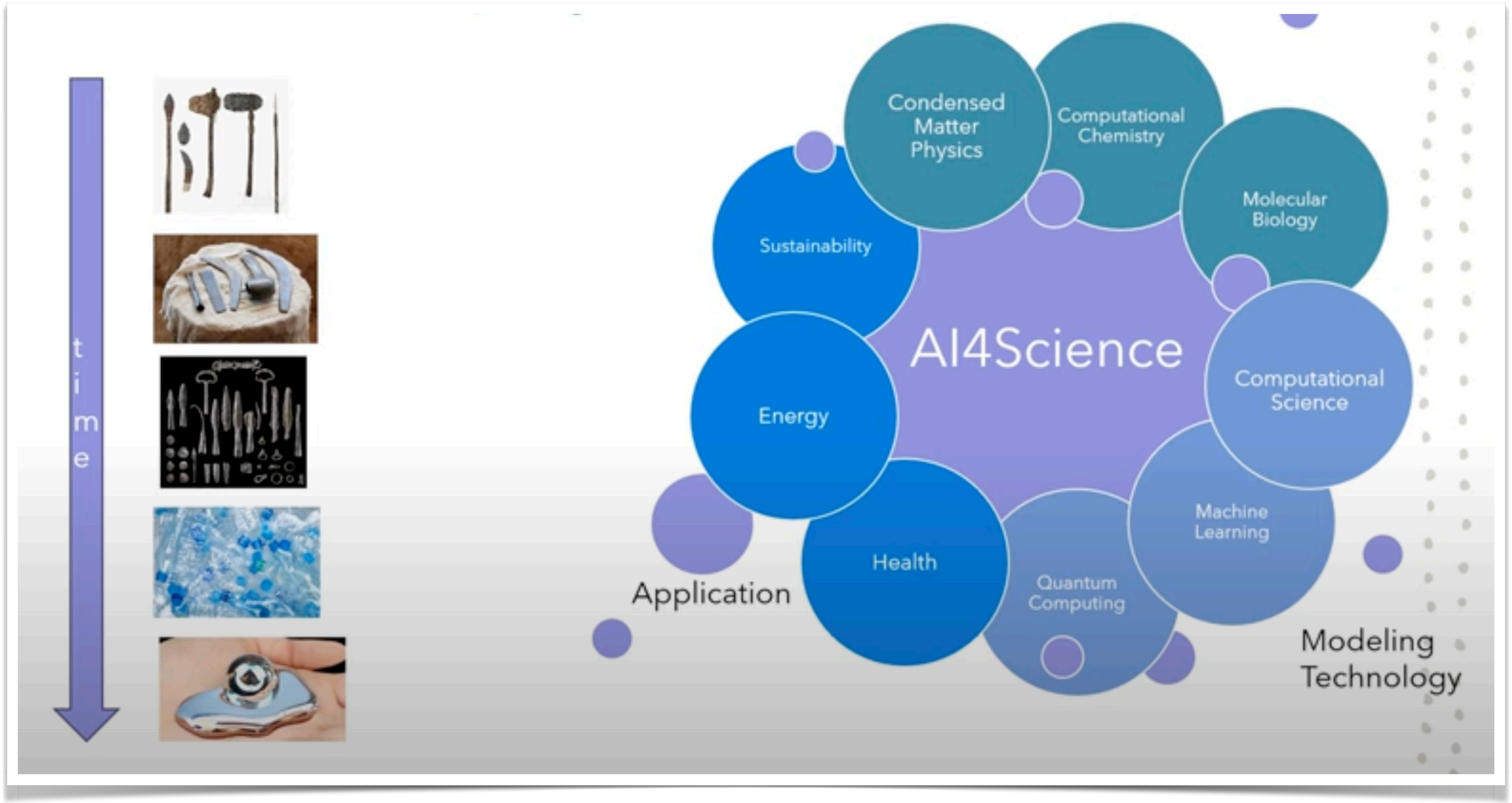


Ben Blaiszik, “2021 AI/ML Publication Statistics and Charts”. Zenodo, Sep. 07, 2022. doi: 10.5281/zenodo.7057437.

Max Welling's talk

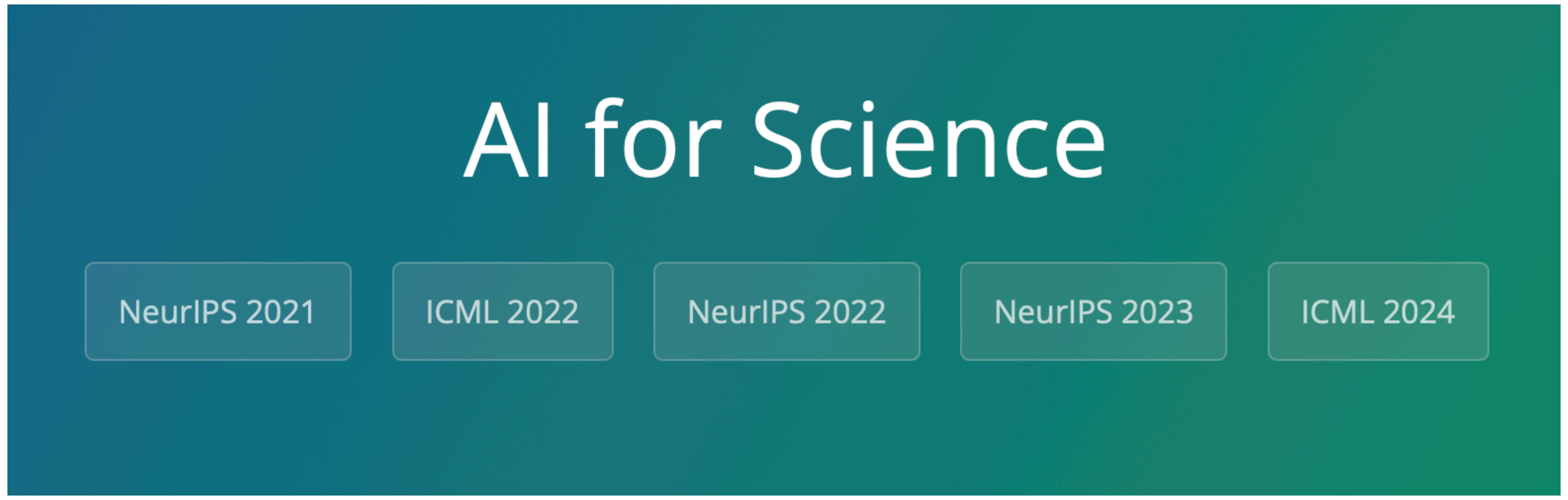
"Where is fundamental physics in AI4Science?" — Lukas Heinrich

- Hint: was not referring to condensed matter physics



AI for Science Workshop Series

https://ai4sciencecommunity.github.io/



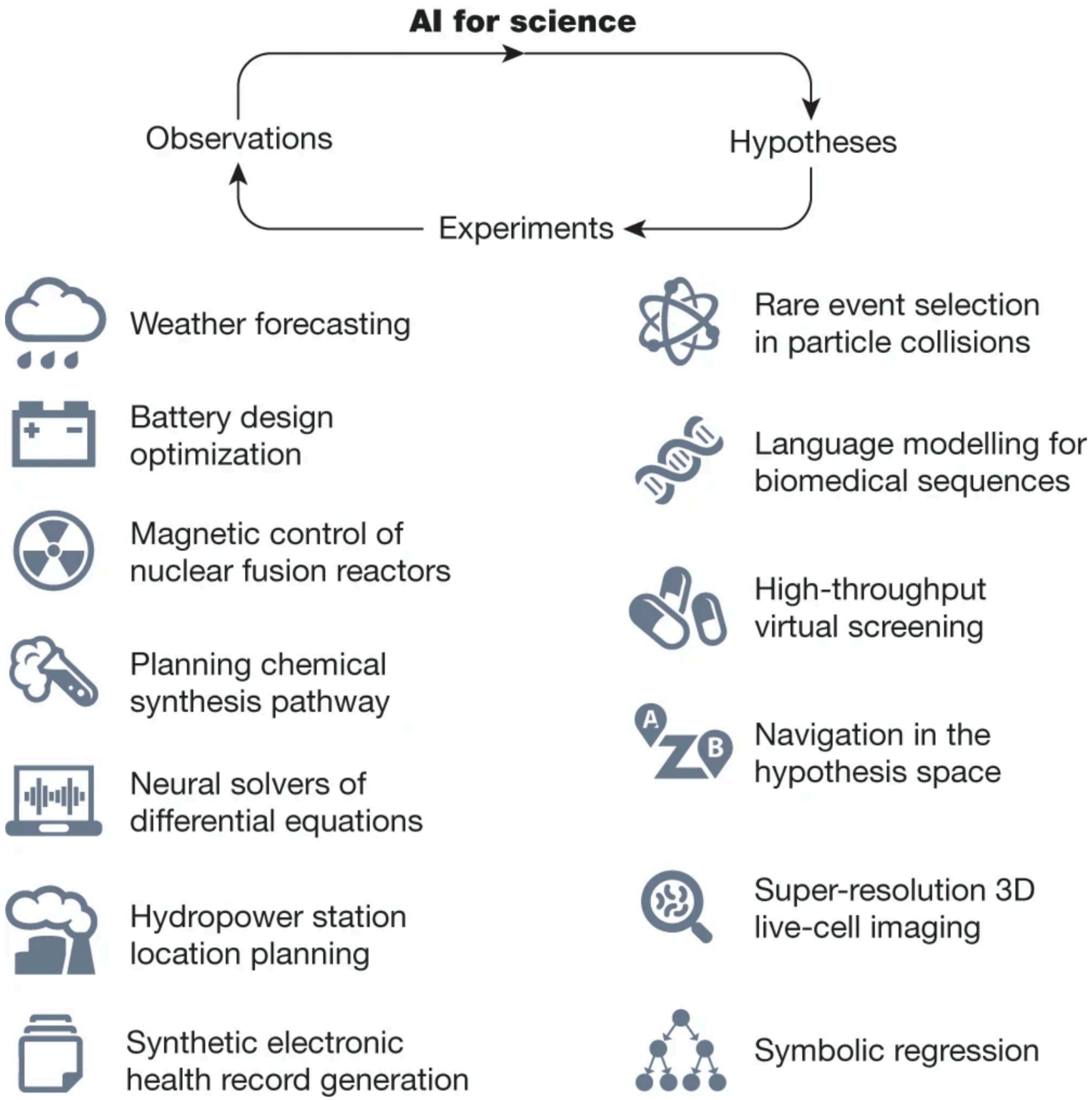
AI for Science

Review | Published: 02 August 2023

Scientific discovery in the age of artificial intelligence

Hanchen Wang, Tianfan Fu, Yuanqi Du, Wenhao Gao, Kexin Huang, Ziming Liu, Payal Chandak, Shengchao Liu, Peter Van Katwyk, Andreea Deac, Anima Anandkumar, Karianne Bergen, Carla P. Gomes, Shirley Ho, Pushmeet Kohli, Joan Lasenby, Jure Leskovec, Tie-Yan Liu, Arjun Manrai, Debora Marks, Bharath Ramsundar, Le Song, Jimeng Sun, Jian Tang, Petar Veličković, Max Welling, Linfeng Zhang, Connor W. Coley, Yoshua Bengio & Marinka Zitnik

Nature **620**, 47–60 (2023) | Cite this article



AI for science

Observations

Hypotheses

Experiments

Weather forecasting

Battery design optimization

Magnetic control of nuclear fusion reactors

Planning chemical synthesis pathway

Neural solvers of differential equations

Hydropower station location planning

Synthetic electronic health record generation

Rare event selection in particle collisions

Language modelling for biomedical sequences

High-throughput virtual screening

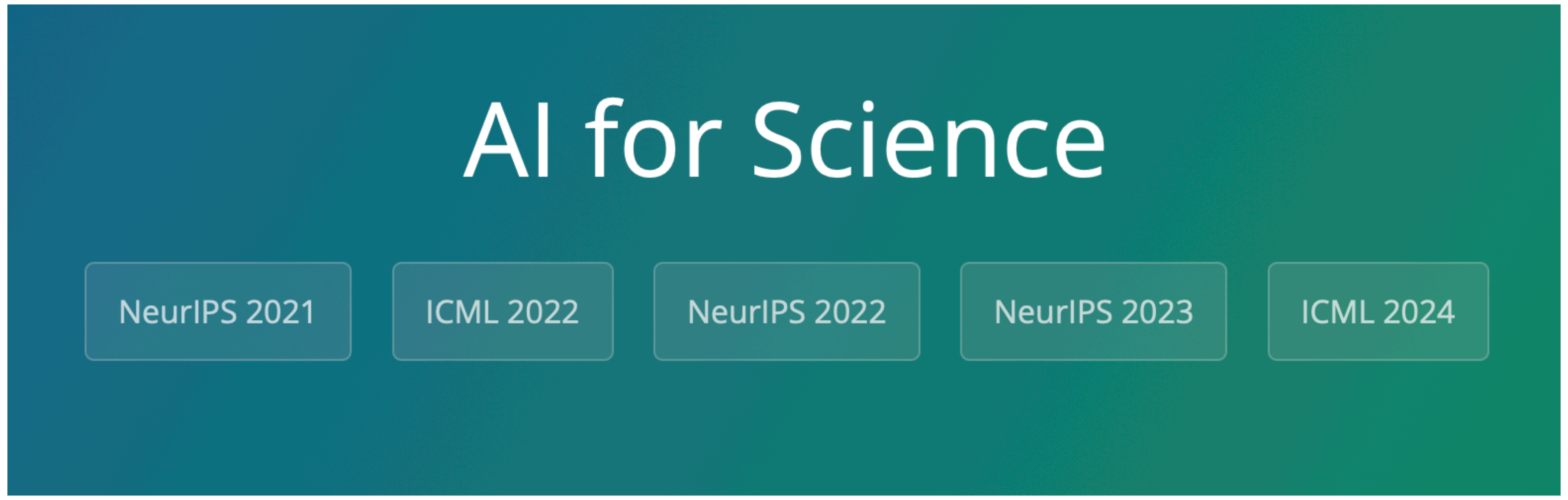
Navigation in the hypothesis space

Super-resolution 3D live-cell imaging

Symbolic regression

AI for Science Workshop Series

https://ai4sciencecommunity.github.io/



Review | Published: 02 August 2023

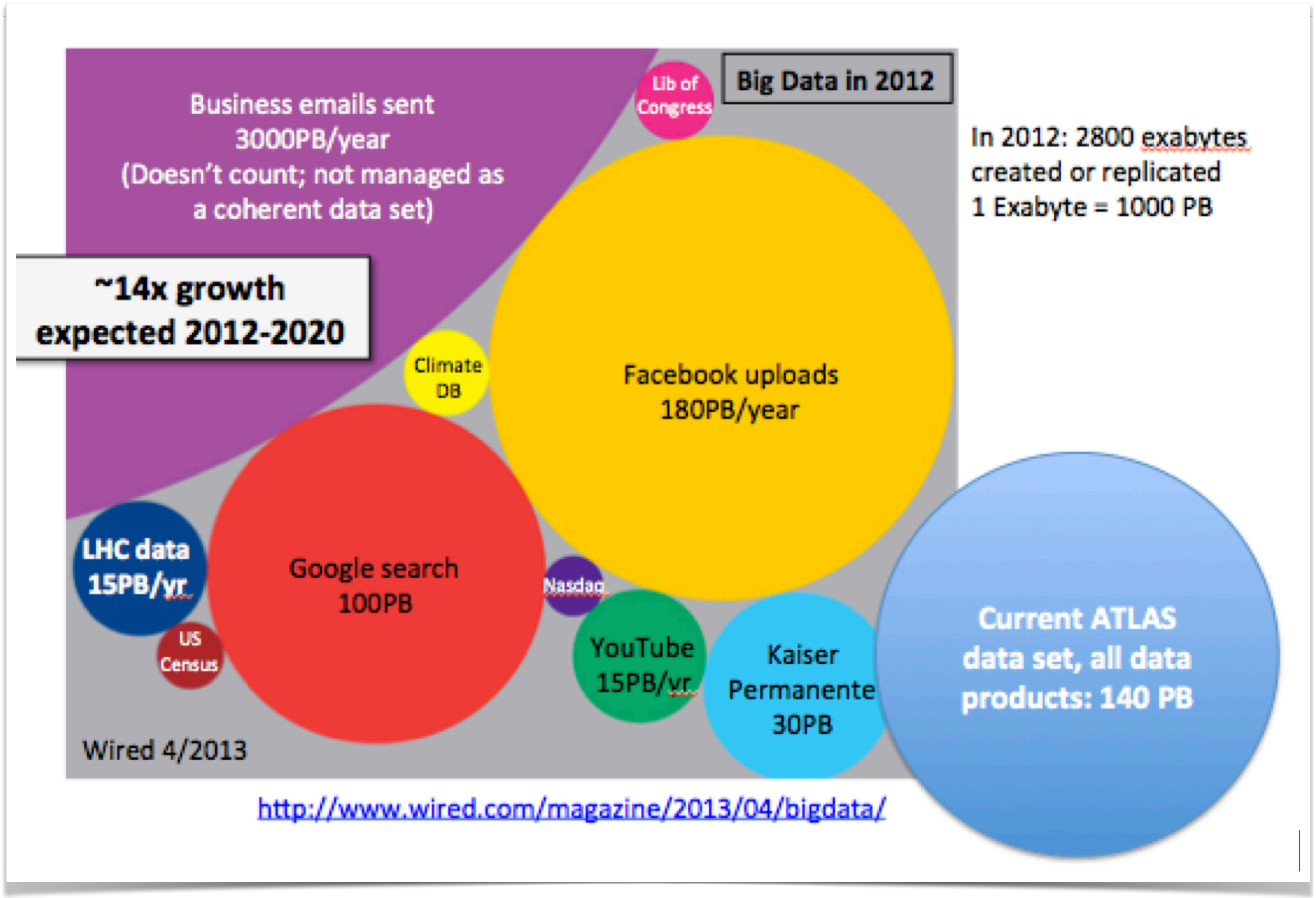
Scientific discovery in the age of artificial intelligence

Hanchen Wang, Tianfan Fu, Yuanqi Du, Wenhao Gao, Kexin Huang, Ziming Liu, Payal Chandak, Shengchao Liu, Peter Van Katwyk, Andreea Deac, Anima Anandkumar, Karianne Bergen, Carla P. Gomes, Shirley Ho, Pushmeet Kohli, Joan Lasenby, Jure Leskovec, Tie-Yan Liu, Arjun Manrai, Debora Marks, Bharath Ramsundar, Le Song, Jimeng Sun, Jian Tang, Petar Veličković, Max Welling, Linfeng Zhang, Connor W. Coley, Yoshua Bengio & Marinka Zitnik

Nature **620**, 47–60 (2023) | Cite this article

Invited Talk	Weinan E	8:15-9:10am
Invited Talk	Shuiwang Ji	9:10-10:05am
Coffee Break	-	10:05-10:15am
Invited Talk	Maria Schuld	10:15-11:10am
Invited Talk	David Baker	11:10-12:05pm
Poster Session		12:05pm-13:00pm
Contributed Talk	DiffDock: Diffusion Steps, Twists, and Turns for Molecular Docking	13:00-13:10pm
Contributed Talk	An "interpretable-by-design" neural network to decipher RNA splicing regulatory logic	13:10-13:20pm
Contributed Talk	FALCON: Fourier Adaptive Learning and Control for Disturbance Rejection Under Extreme Turbulence	13:20-13:30pm
Contributed Talk	Structural Causal Model for Molecular Dynamics Simulation	13:30-13:40pm
Contributed Talk	Structure-Inducing Pre-training	13:40-13:50pm
Panel Discussion	Panelists: Peter Lee, Sara-Jane Dunn, Arvind Rajpal, Alexander Rives, Matt Welborn Moderator: Max Welling	13:55-14:55pm
Coffee Break	-	14:55-15:10pm
Invited Talk	Jimeng Sun	15:10-16:05pm
Invited Talk	Tess Smidt	16:05-17:00pm

Compare with **Big Data**



In 2012: 2800 exabytes created or replicated
1 Exabyte = 1000 PB

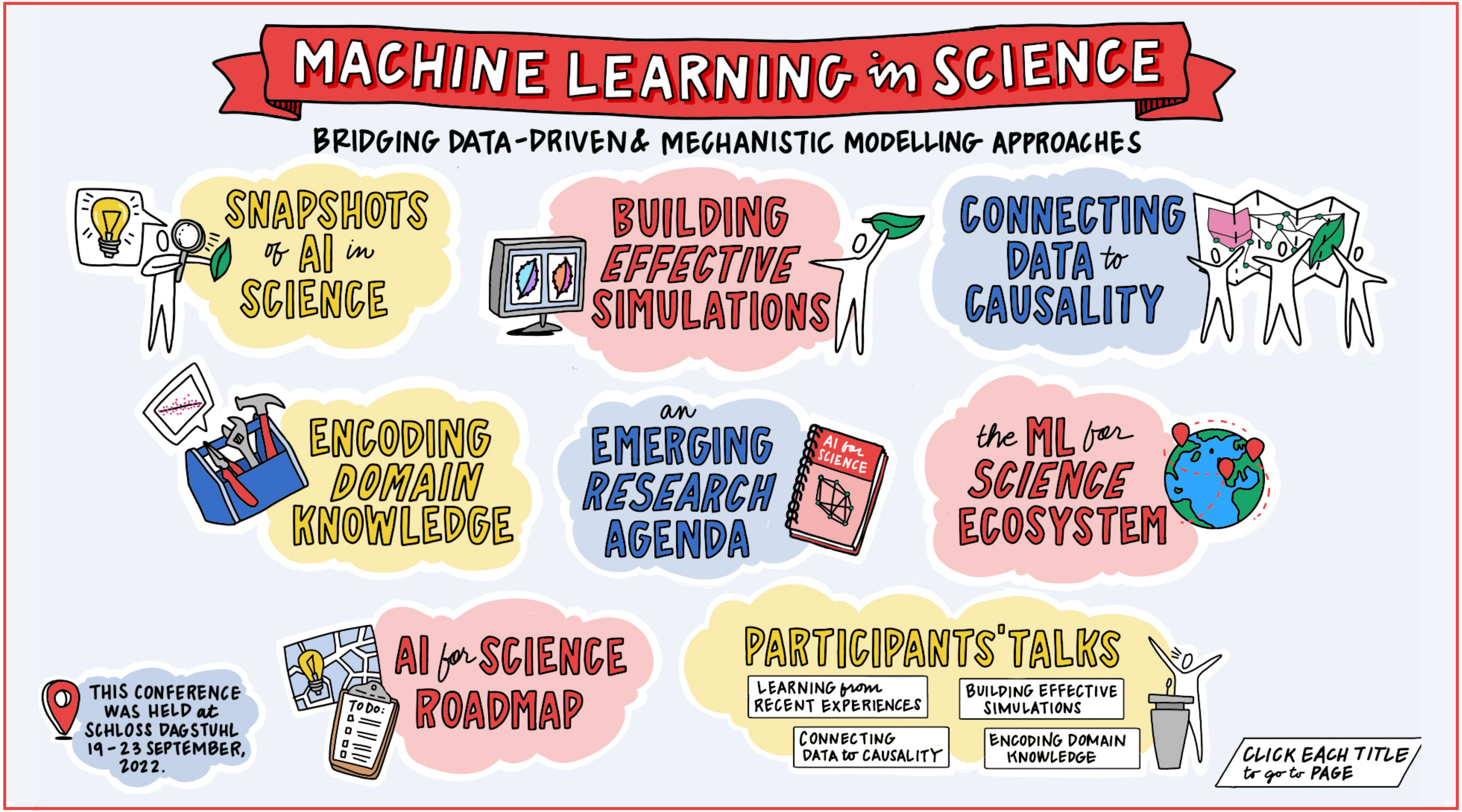
Wired 4/2013

http://www.wired.com/magazine/2013/04/bigdata/

What are some of the dominant narratives about AI for Science?

How does fundamental physics fit in?

https://acceleratescience.github.io/resources/machine-learning-for-science-bridging-data-driven-and-mechanistic-modelling.html



MACHINE LEARNING in SCIENCE

BRIDGING DATA-DRIVEN & MECHANISTIC MODELLING APPROACHES

SNAPSHOTS of AI in SCIENCE

BUILDING EFFECTIVE SIMULATIONS

CONNECTING DATA to CAUSALITY

ENCODING DOMAIN KNOWLEDGE

an EMERGING RESEARCH AGENDA

the ML for SCIENCE ECOSYSTEM

THIS CONFERENCE WAS HELD at SCHLOSS DAGSTUHL 19-23 SEPTEMBER, 2022.

AI for SCIENCE ROADMAP

PARTICIPANTS' TALKS

- LEARNING from RECENT EXPERIENCES
- BUILDING EFFECTIVE SIMULATIONS
- CONNECTING DATA to CAUSALITY
- ENCODING DOMAIN KNOWLEDGE

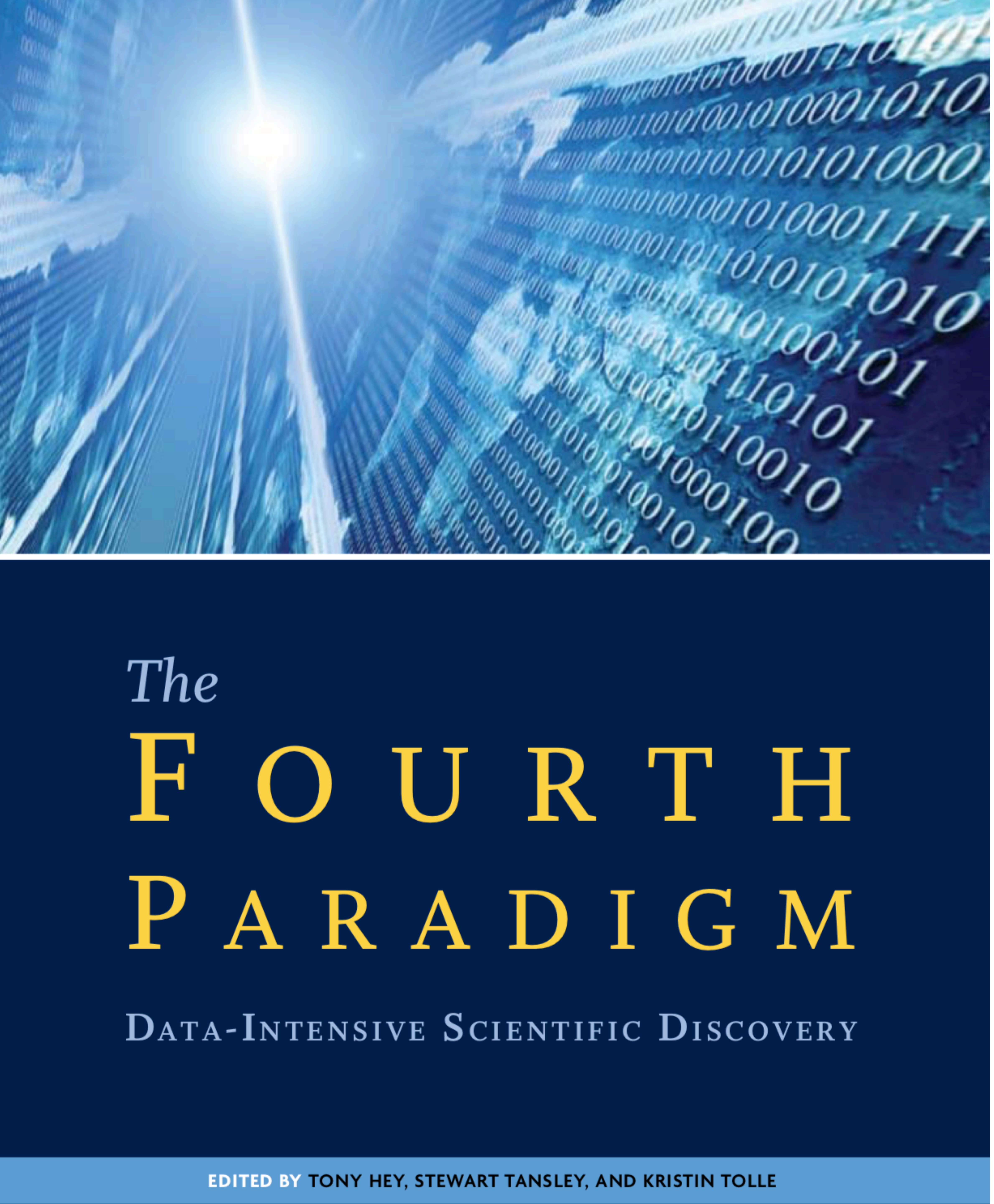
CLICK EACH TITLE to go to PAGE

UNIVERSITY OF CAMBRIDGE

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

NYU

WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON



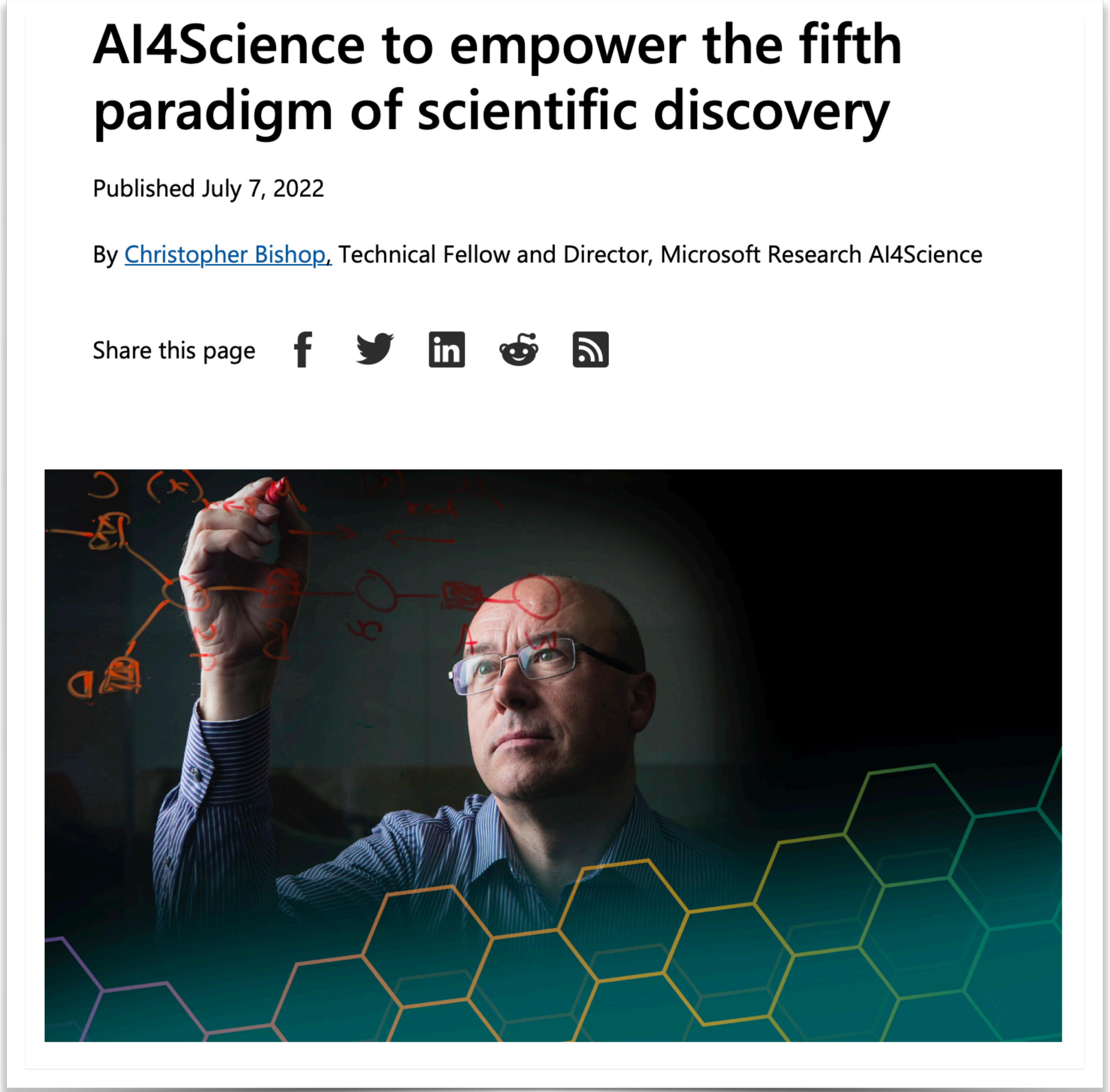
The
FOURTH
PARADIGM

DATA-INTENSIVE SCIENTIFIC DISCOVERY

EDITED BY TONY HEY, STEWART TANSLEY, AND KRISTIN TOLLE

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

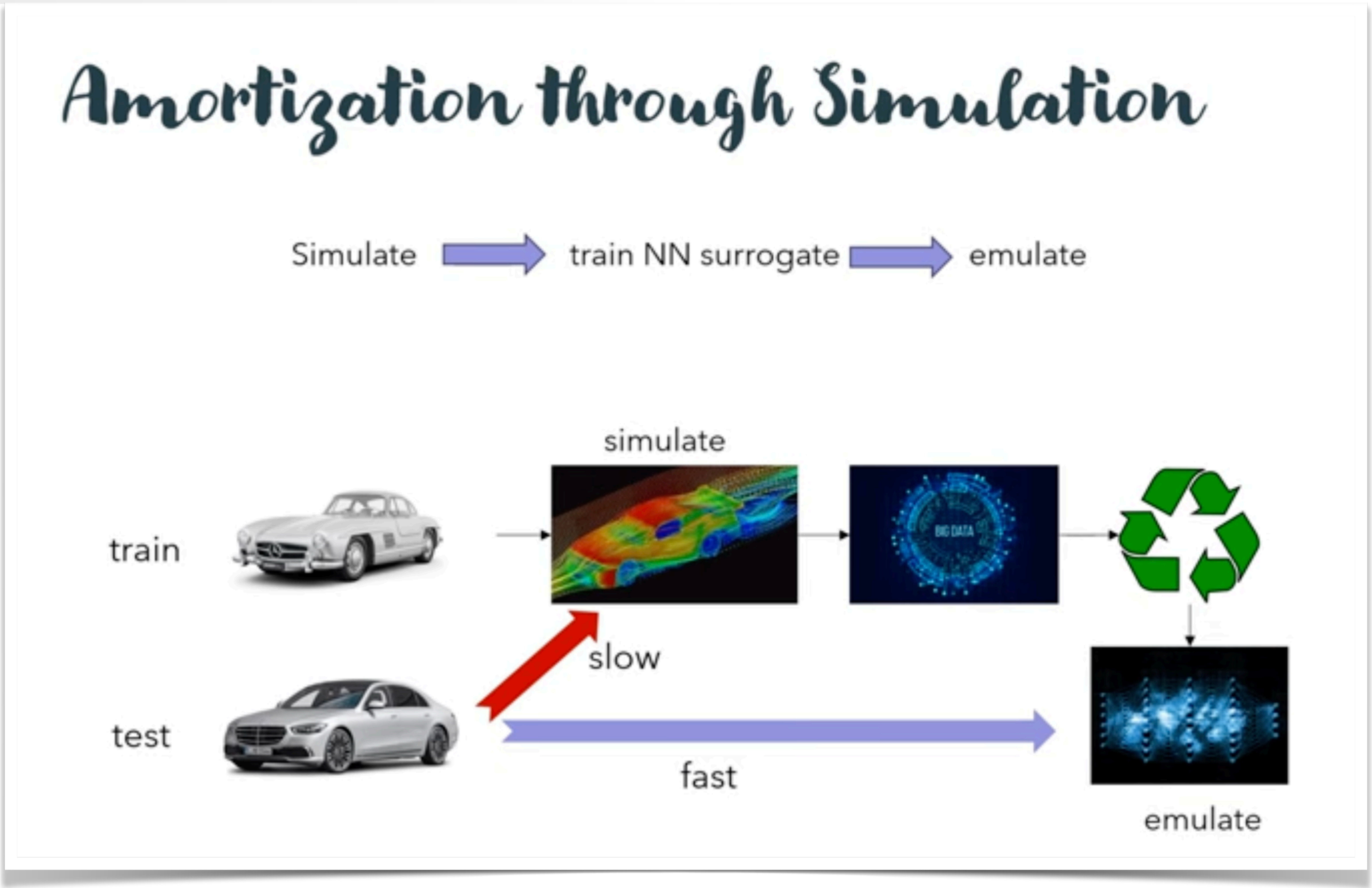
By Christopher Bishop, Technical Fellow and Director, Microsoft Research AI4Science

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

"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

Simulation & Emulators



Amortization through Simulation

Simulate → train NN surrogate → emulate

simulate

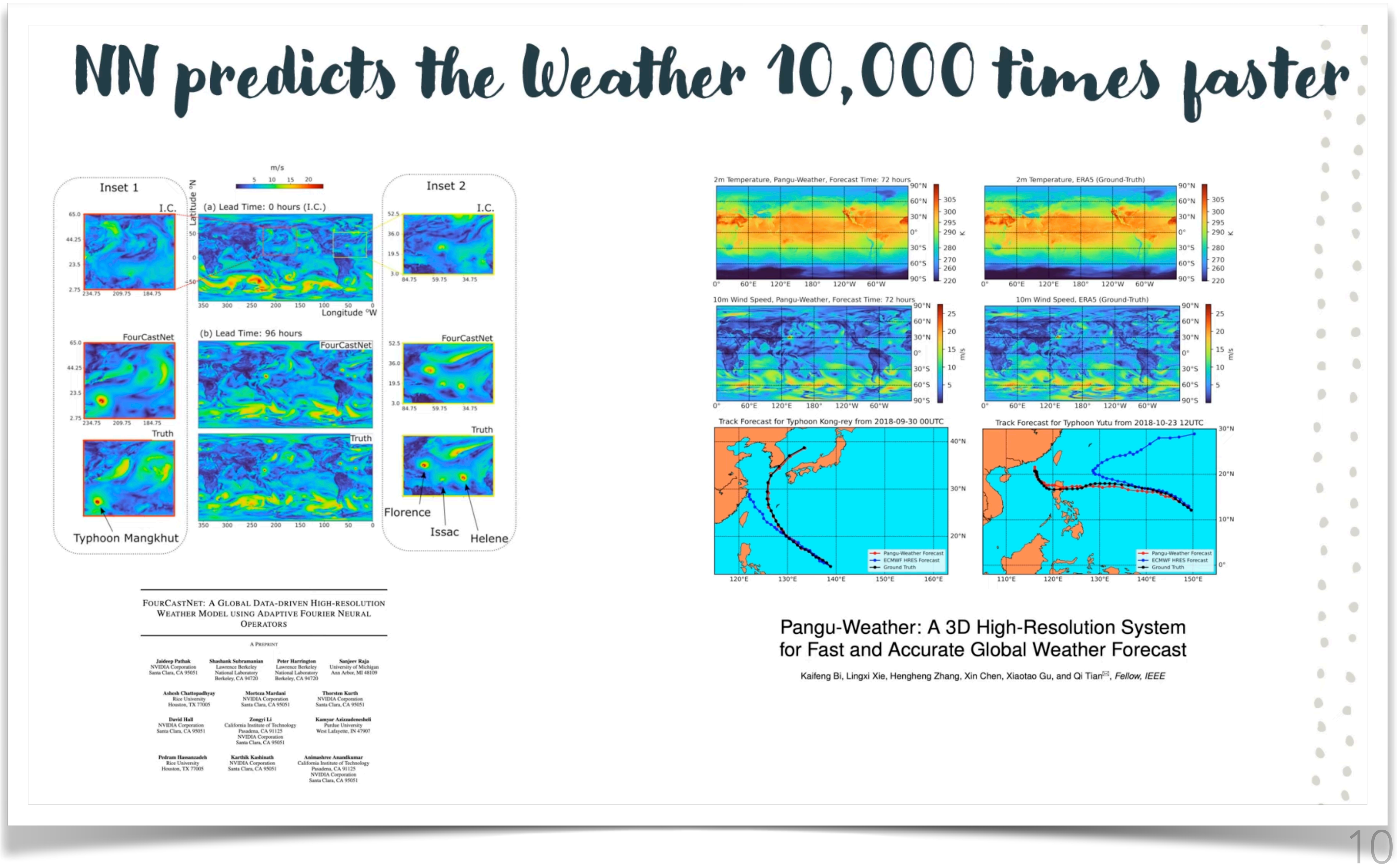
train

test

slow

fast

emulate



NN predicts the Weather 10,000 times faster

Inset 1

Inset 2

I.C.

(a) Lead Time: 0 hours (I.C.)

(b) Lead Time: 96 hours

FourCastNet

Truth

Typhoon Mangkhut

Florence

Issac

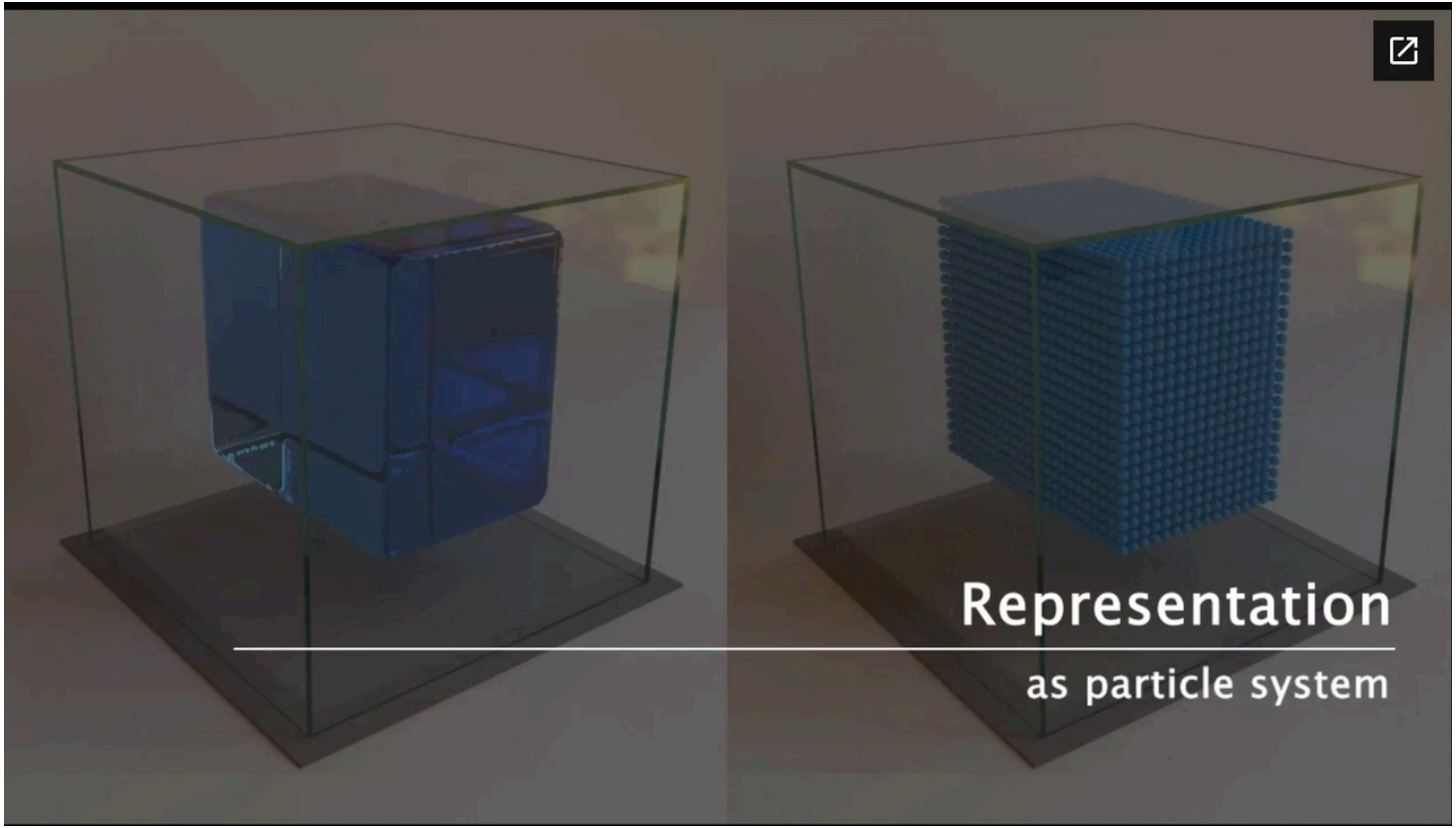
Helene

FourCastNet: A Global Data-driven High-resolution Weather Model using Adaptive Fourier Neural Operators

Pangu-Weather: A 3D High-Resolution System for Fast and Accurate Global Weather Forecast

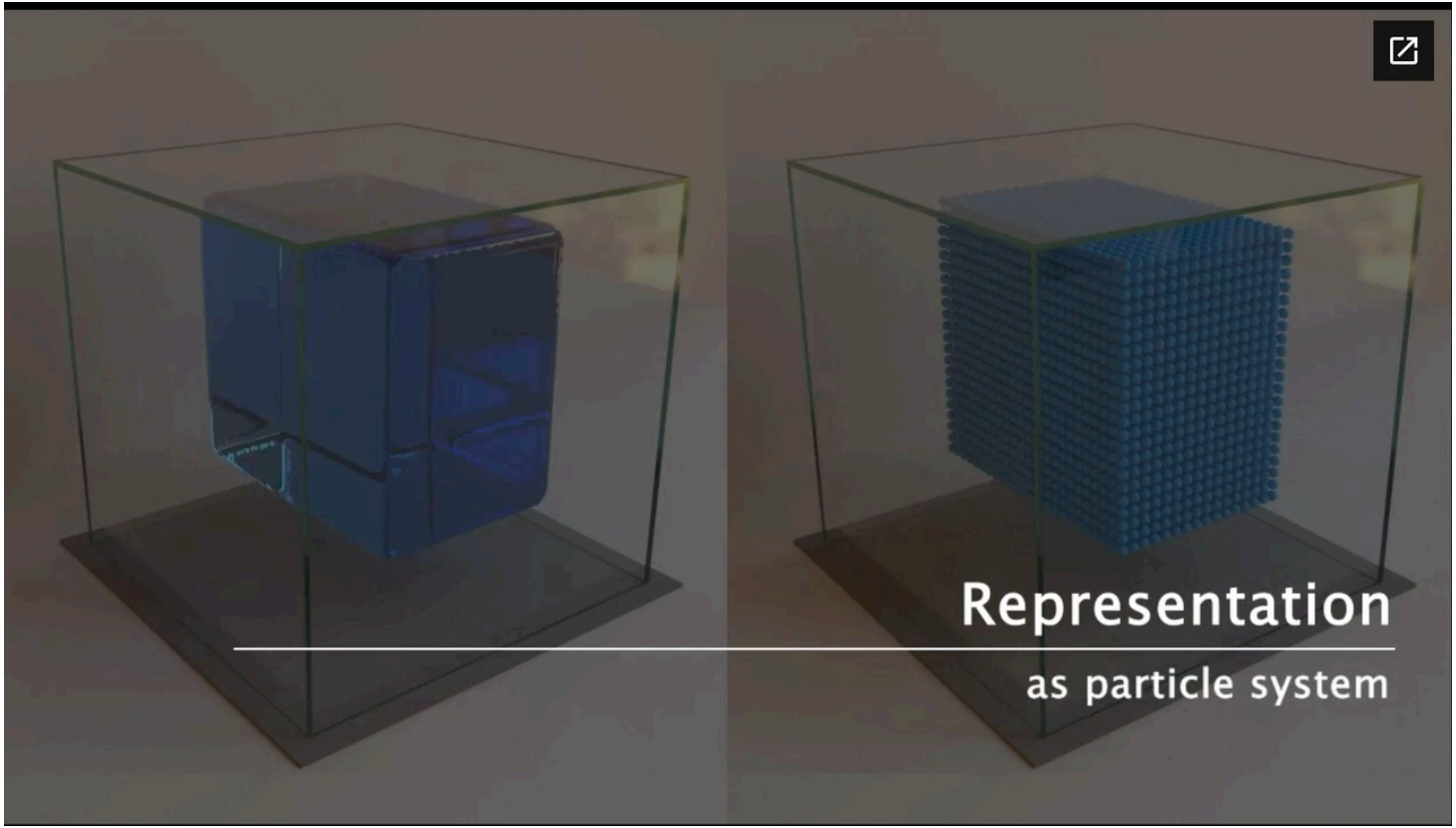
Kaifeng Bi, Lingxi Xie, Hengheng Zhang, Xin Chen, Xiaotao Gu, and Qi Tian, *Fellow, IEEE*

AI/ML as emulators of complex simulations



Alvaro Sanchez-Gonzalez, Jonathan Godwin, Tobias Pfaff, Rex Ying, Jure Leskovec, Peter W Battaglia, arXiv:2002.09405

AI/ML as emulators of complex simulations



Alvaro Sanchez-Gonzalez, Jonathan Godwin, Tobias Pfaff, Rex Ying, Jure Leskovec, Peter W Battaglia, arXiv:2002.09405

In fusion energy

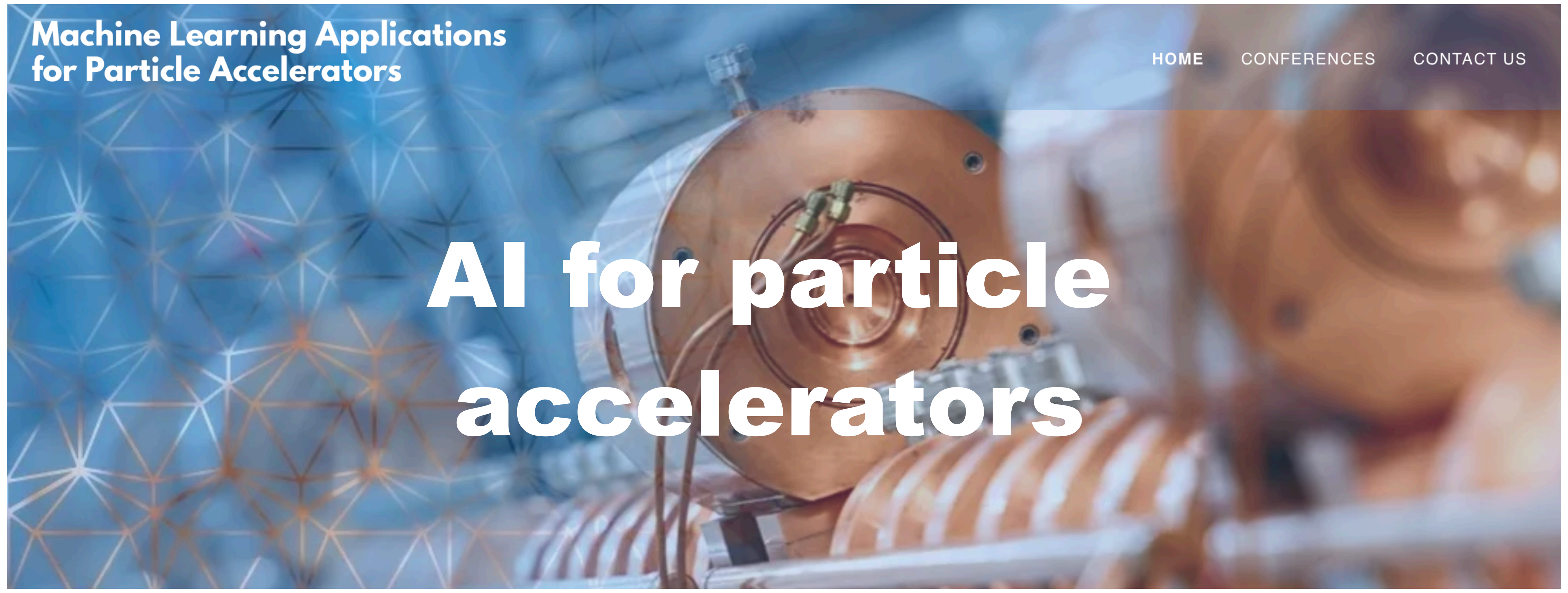
Research

Accelerating fusion science through learned plasma control

February 16, 2022

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

Verena Kain's talk



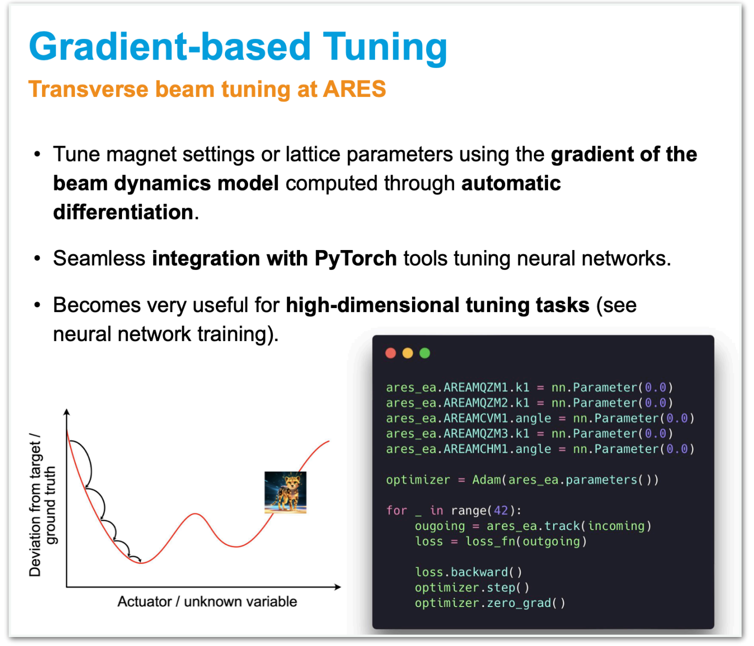
AI for particle accelerators

Differentiable simulation codes

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

Cheetah – A High-speed Differentiable Beam Dynamics Simulation for Machine Learning Applications

4th ICFA Machine Learning Workshop



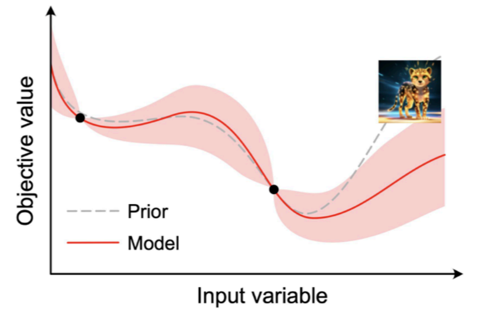
Gradient-based Tuning

Transverse beam tuning at ARES

- Tune magnet settings or lattice parameters using the **gradient of the beam dynamics model** computed through **automatic differentiation**.
- Seamless **integration with PyTorch** tools tuning neural networks.
- Becomes very useful for **high-dimensional tuning tasks** (see neural network training).

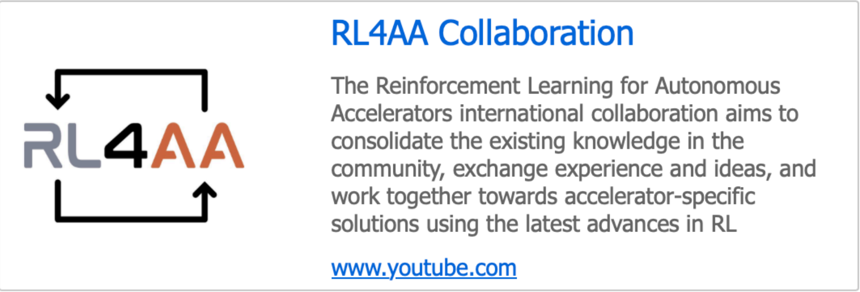
```
ares_ea.AREAMQZM1.k1 = nn.Parameter(0.0)
ares_ea.AREAMQZM2.k1 = nn.Parameter(0.0)
ares_ea.AREAMCVM1.angle = nn.Parameter(0.0)
ares_ea.AREAMQZM3.k1 = nn.Parameter(0.0)
ares_ea.AREAMCHM1.angle = nn.Parameter(0.0)

optimizer = Adam(ares_ea.parameters())

for _ in range(42):
    ougoing = ares_ea.track(incoming)
    loss = loss_fn(outgoing)

    loss.backward()
    optimizer.step()
    optimizer.zero_grad()
```

- 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.



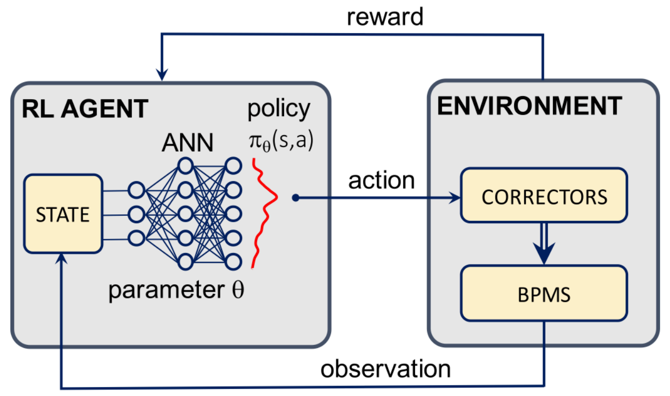
RL4AA Collaboration

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 <u>for</u> 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!

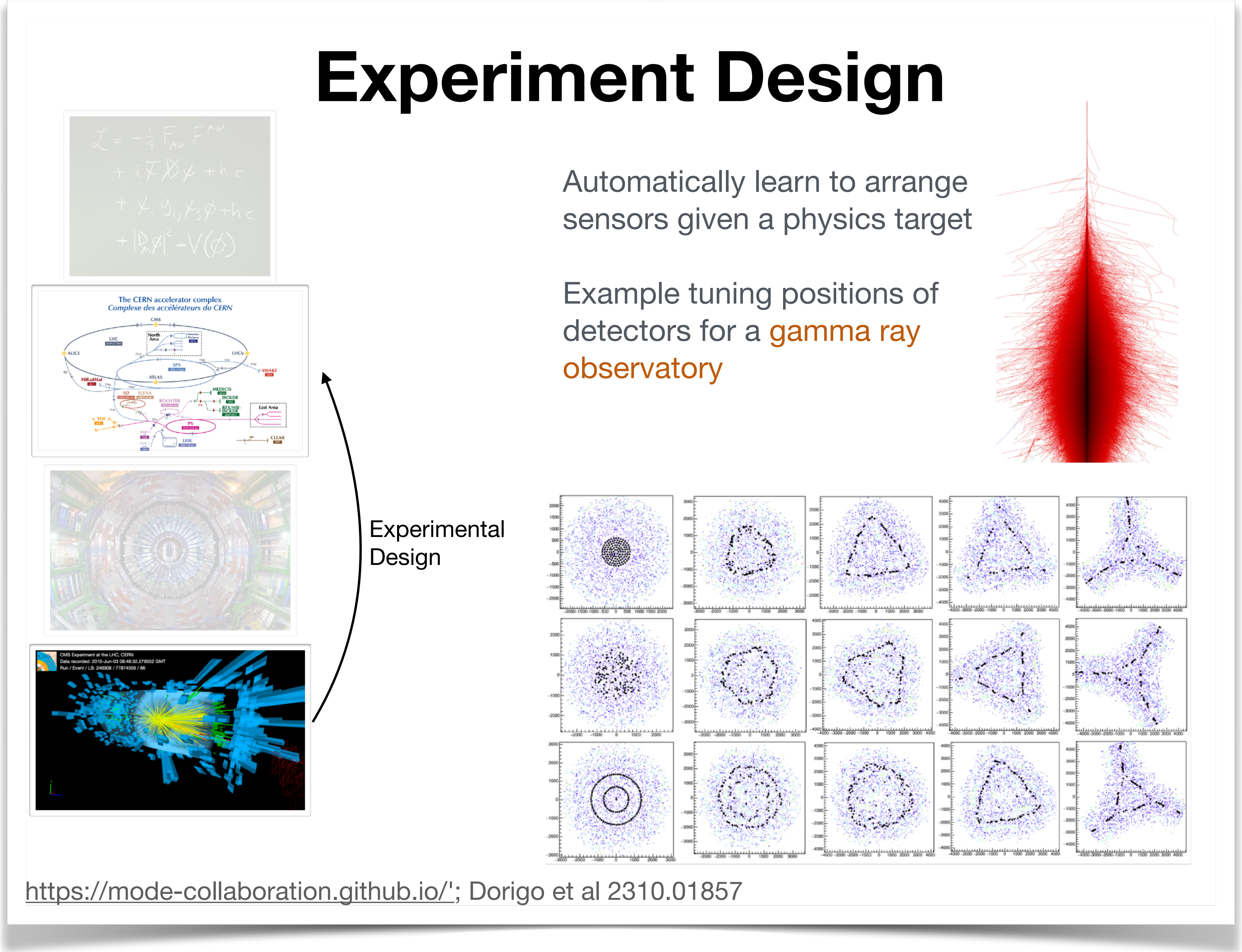
- Not sample-efficient enough
- Safety constraints

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

Surrogates and differentiable programming

Third MODE Workshop on
Differentiable Programming for Experiment Design

Princeton University
24-26 July, 2023

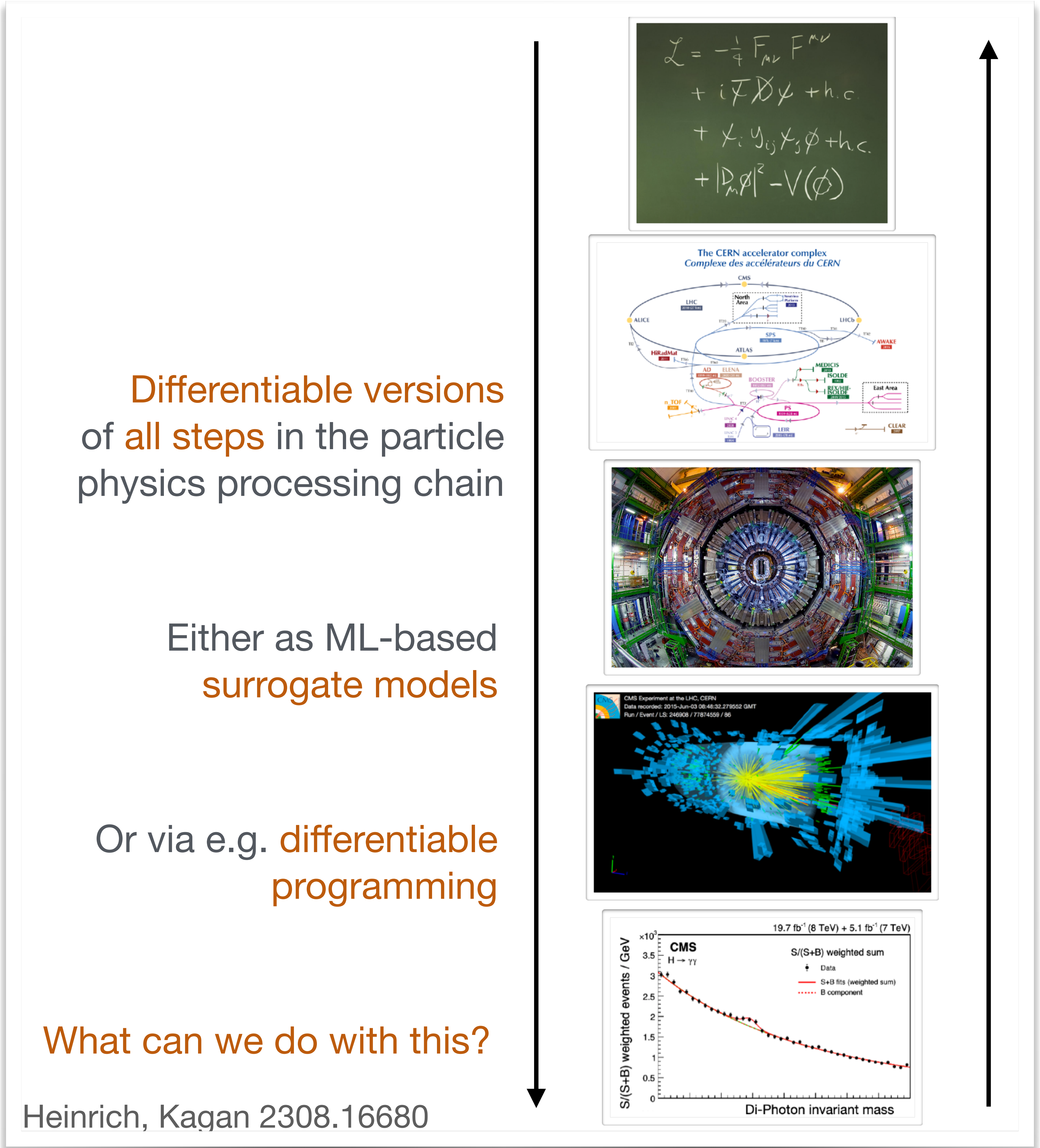


Experiment Design

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

Slides from Gregor Kasieczka's talk

Inductive Bias

Compositionality

Relationships

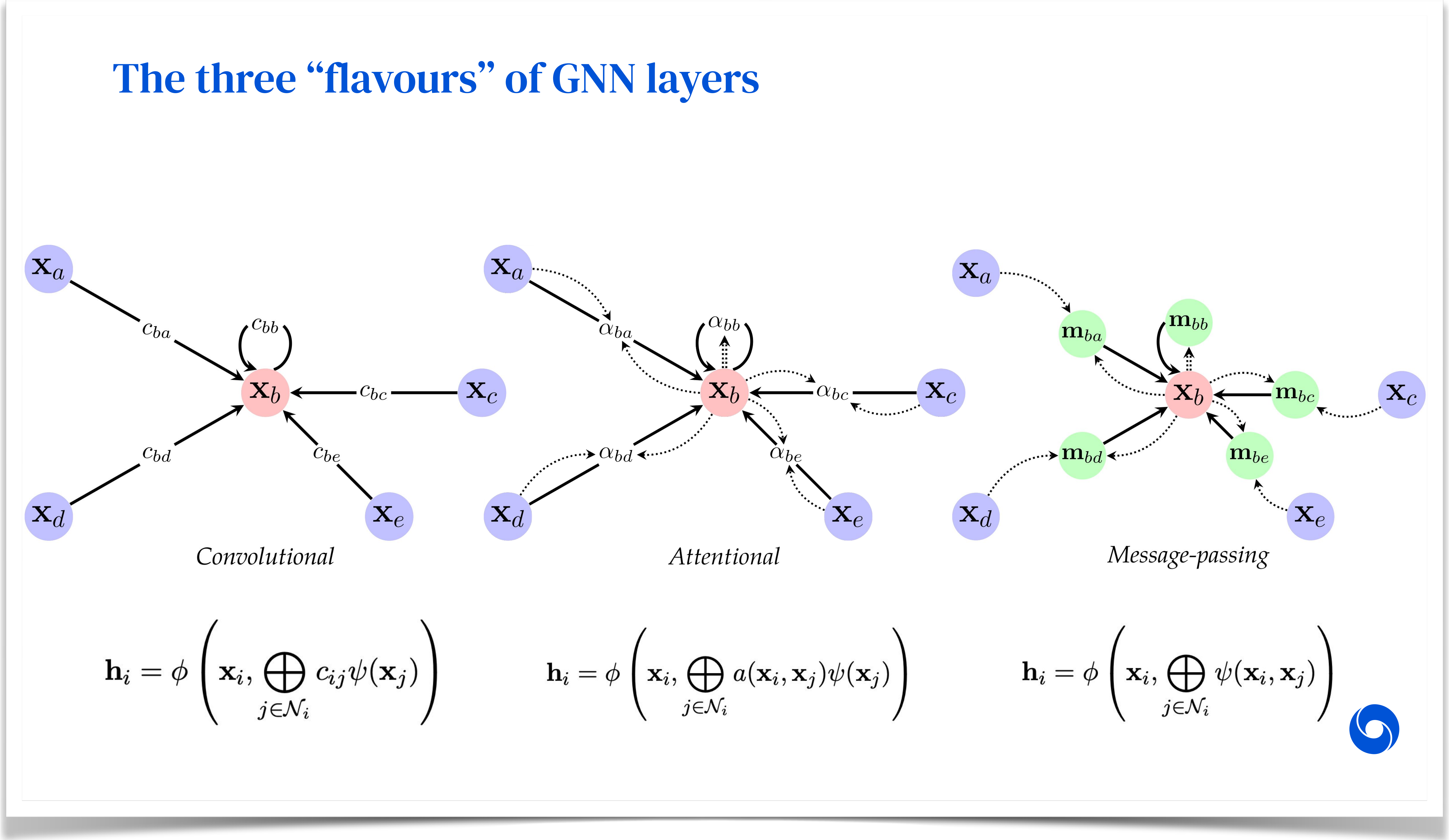
Symmetry

Causality

Geometric Deep Learning

A hot topic in AI/ML research

- How do we incorporate domain knowledge?
- Should we? "The bitter lesson"



The three "flavours" of GNN layers

Convolutional

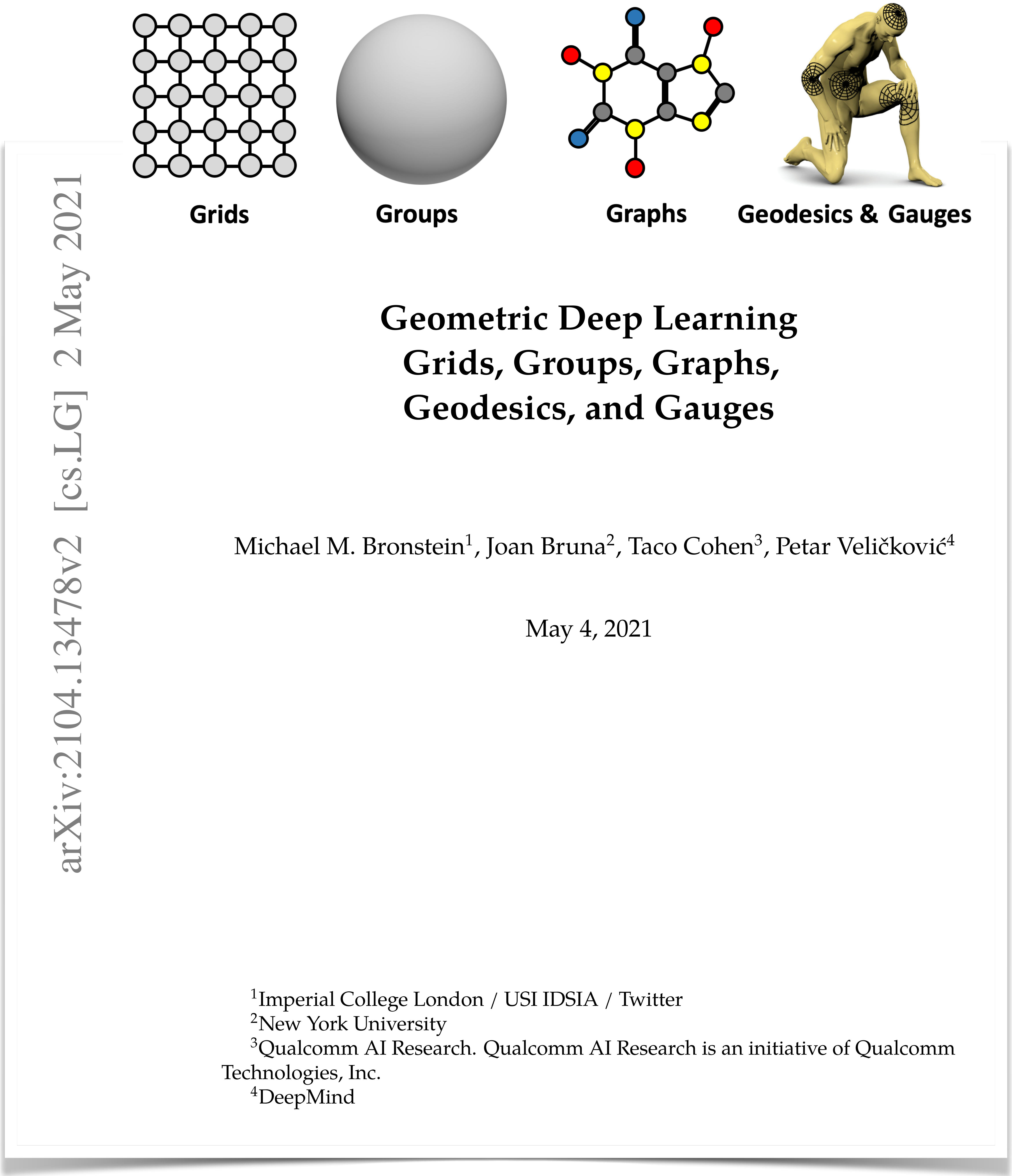
Attentional

Message-passing

$$\mathbf{h}_i = \phi\left(\mathbf{x}_i, \bigoplus_{j\in\mathcal{N}_i} c_{ij}\psi(\mathbf{x}_j)\right)$$

$$\mathbf{h}_i = \phi\left(\mathbf{x}_i, \bigoplus_{j\in\mathcal{N}_i} a(\mathbf{x}_i, \mathbf{x}_j)\psi(\mathbf{x}_j)\right)$$

$$\mathbf{h}_i = \phi\left(\mathbf{x}_i, \bigoplus_{j\in\mathcal{N}_i} \psi(\mathbf{x}_i, \mathbf{x}_j)\right)$$



arXiv:2104.13478v2 [cs.LG] 2 May 2021

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

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

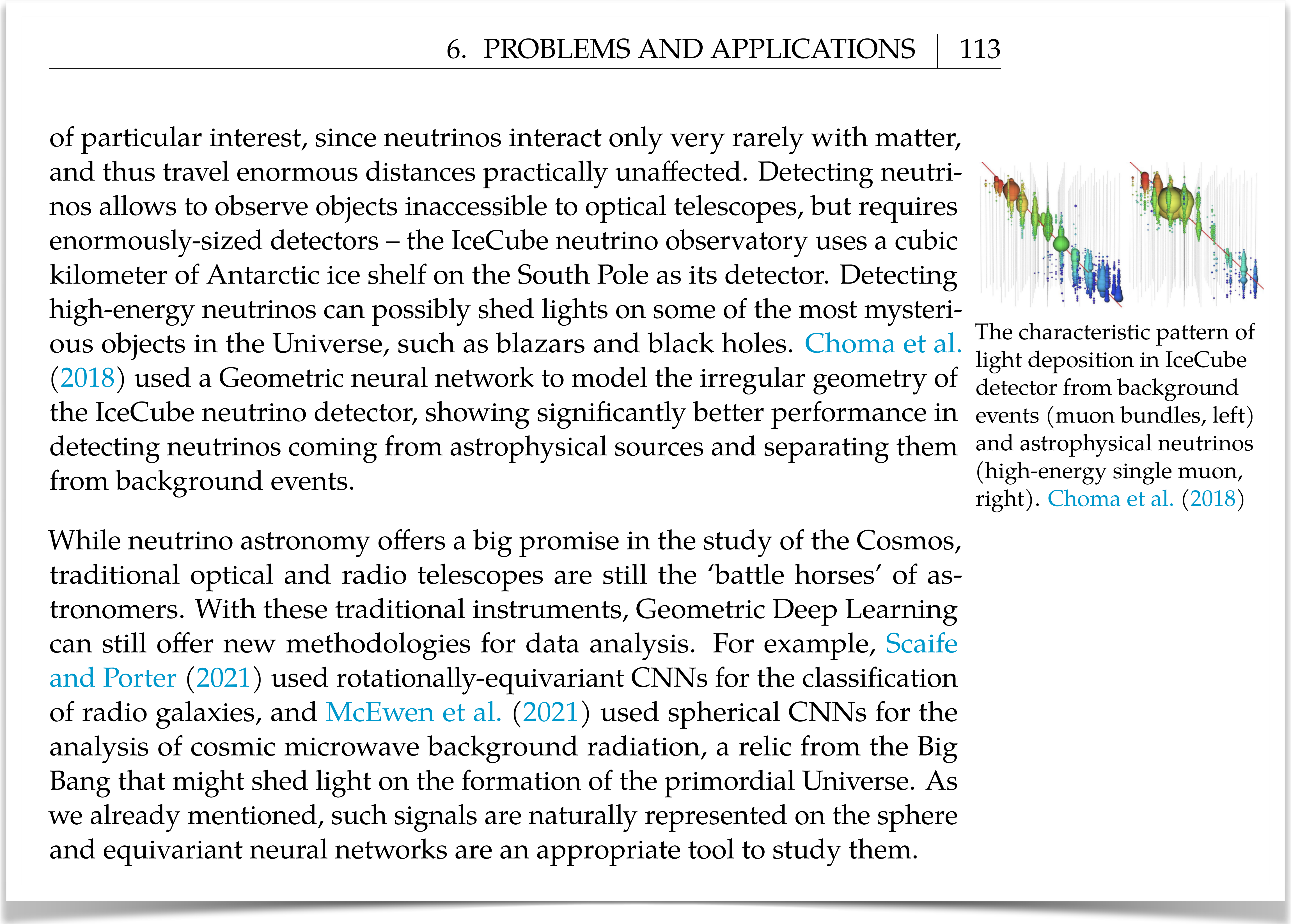
May 4, 2021

[1] Imperial College London / USI IDSIA / Twitter
[2] New York University
[3] Qualcomm AI Research. Qualcomm AI Research is an initiative of Qualcomm Technologies, Inc.
[4] DeepMind

Geometric Deep Learning

Particle physics & astrophysics are highlighted in review

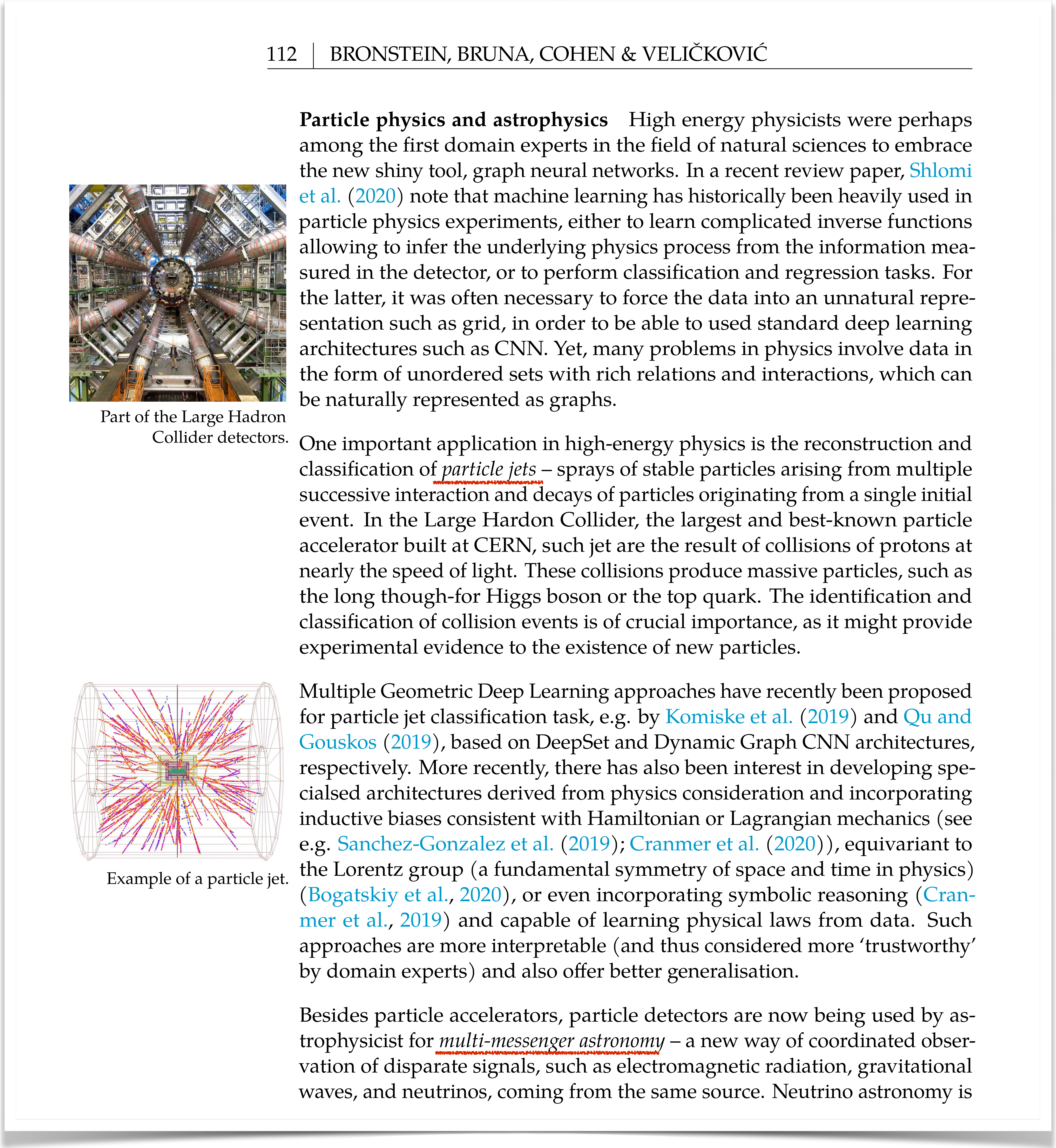
- Continues to be an effective connection with AI community



of particular interest, since neutrinos interact only very rarely with matter, and thus travel enormous distances practically unaffected. Detecting neutrinos allows to observe objects inaccessible to optical telescopes, but requires enormously-sized detectors – the IceCube neutrino observatory uses a cubic kilometer of Antarctic ice shelf on the South Pole as its detector. Detecting high-energy neutrinos can possibly shed lights on some of the most mysterious objects in the Universe, such as blazars and black holes. Choma et al. (2018) used a Geometric neural network to model the irregular geometry of the IceCube neutrino detector, showing significantly better performance in detecting neutrinos coming from astrophysical sources and separating them from background events.

The characteristic pattern of light deposition in IceCube detector from background events (muon bundles, left) and astrophysical neutrinos (high-energy single muon, right). Choma et al. (2018)

While neutrino astronomy offers a big promise in the study of the Cosmos, traditional optical and radio telescopes are still the 'battle horses' of astronomers. With these traditional instruments, Geometric Deep Learning can still offer new methodologies for data analysis. For example, Scaife and Porter (2021) used rotationally-equivariant CNNs for the classification of radio galaxies, and McEwen et al. (2021) used spherical CNNs for the analysis of cosmic microwave background radiation, a relic from the Big Bang that might shed light on the formation of the primordial Universe. As we already mentioned, such signals are naturally represented on the sphere and equivariant neural networks are an appropriate tool to study them.



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.

Part of the Large Hadron Collider detectors.

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.

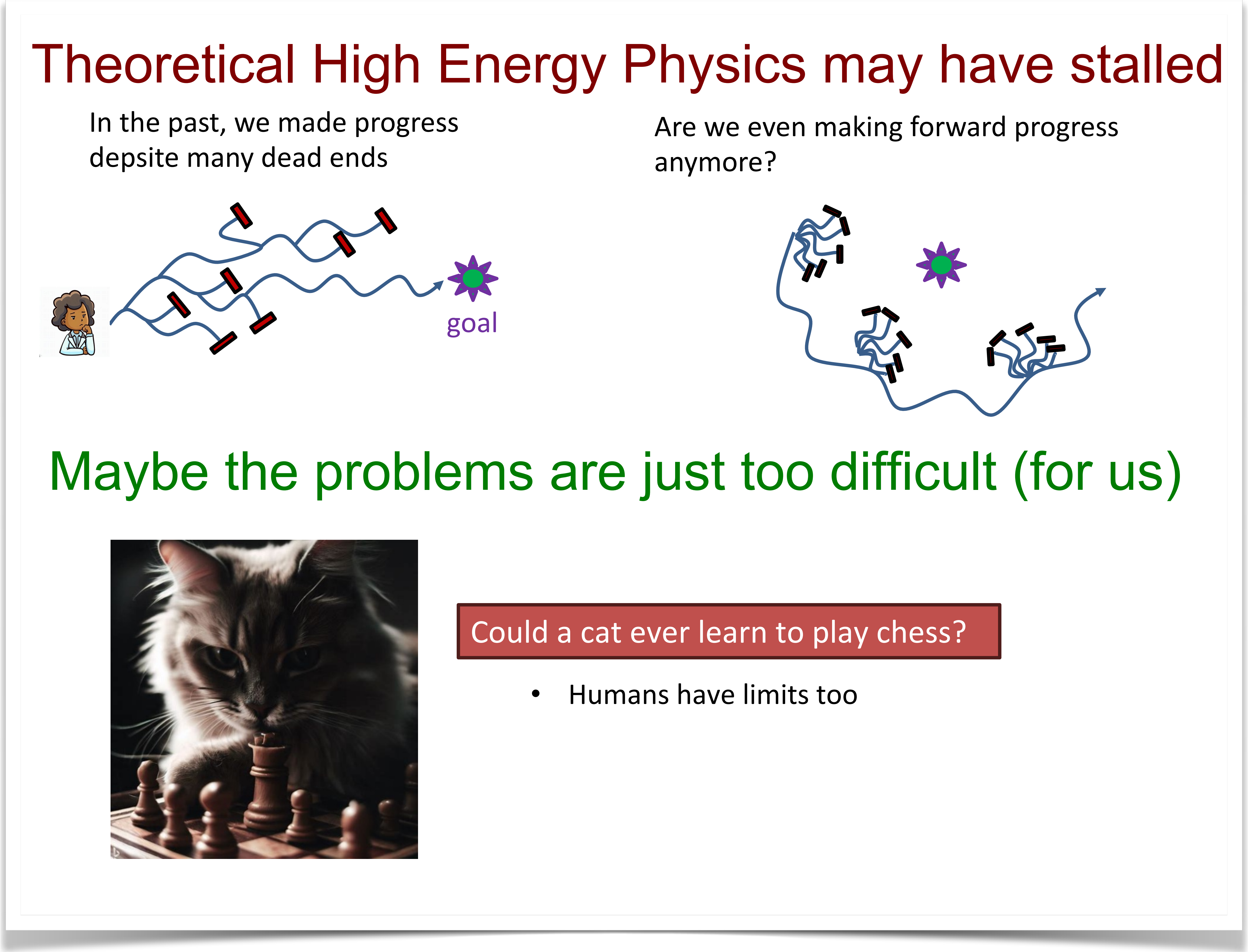
Example of a particle jet.

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

Scientific Understanding

Matt Schwartz's talk



Theoretical High Energy Physics may have stalled

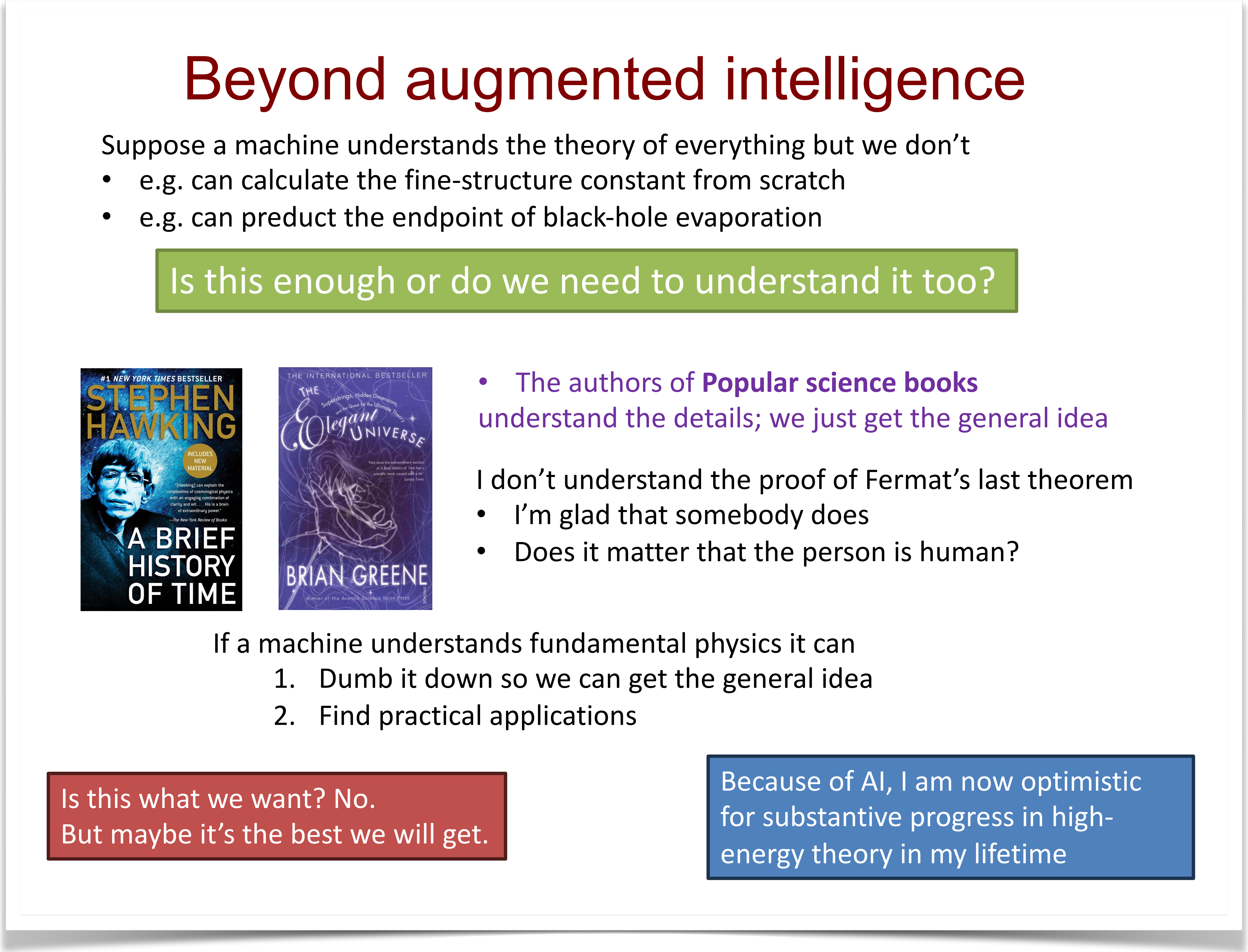
In the past, we made progress depsite many dead ends

Are we even making forward progress anymore?

Maybe the problems are just too difficult (for us)

Could a cat ever learn to play chess?

- Humans have limits too



Beyond augmented intelligence

Suppose a machine understands the theory of everything but we don't
- e.g. can calculate the fine-structure constant from scratch
- e.g. can preduct the endpoint of black-hole evaporation

Is this enough or do we need to understand it too?

- The authors of **Popular science books** understand the details; we just get the general idea

I don't understand the proof of Fermat's last theorem
- I'm glad that somebody does
- Does it matter that the person is human?

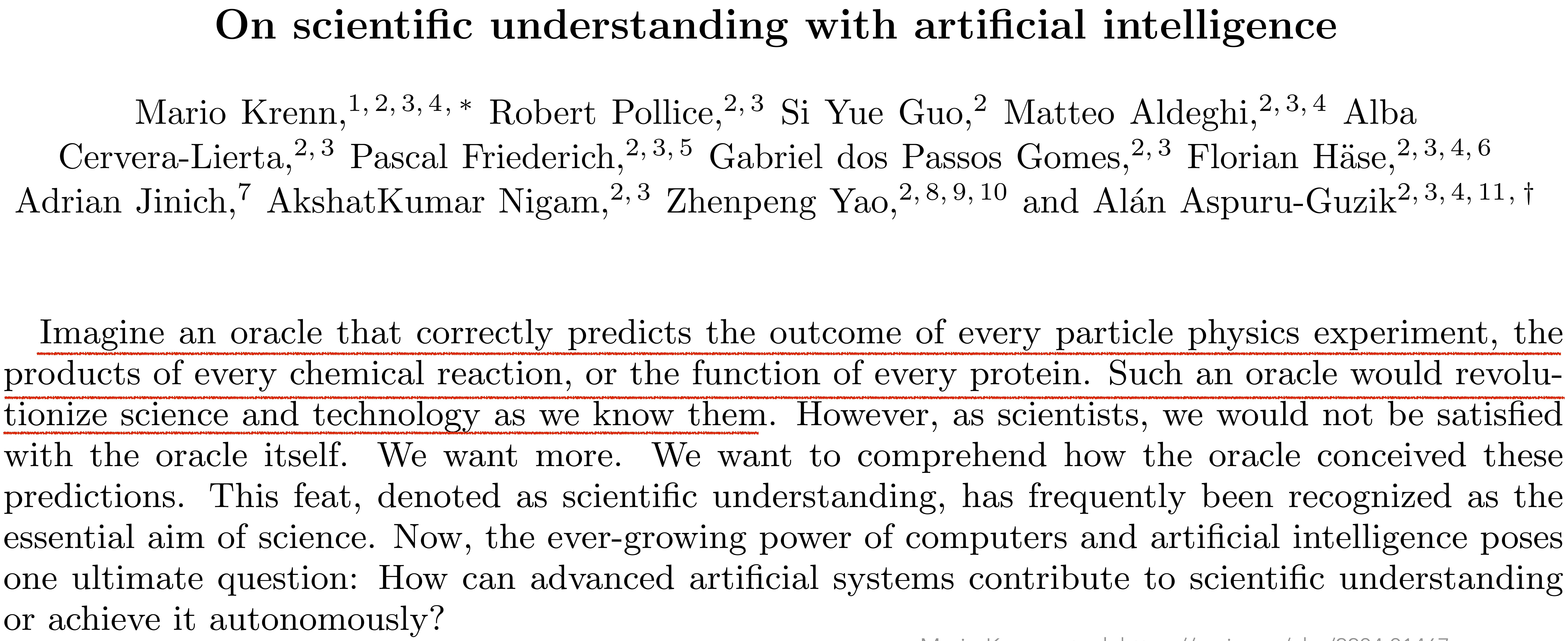
If a machine understands fundamental physics it can
1. Dumb it down so we can get the general idea
2. Find practical applications

Is this what we want? No.
But maybe it's the best we will get.

Because of AI, I am now optimistic for substantive progress in high-energy theory in my lifetime

Scientific understanding

How does AI enable or enhance scientific understanding?



On scientific understanding with artificial intelligence

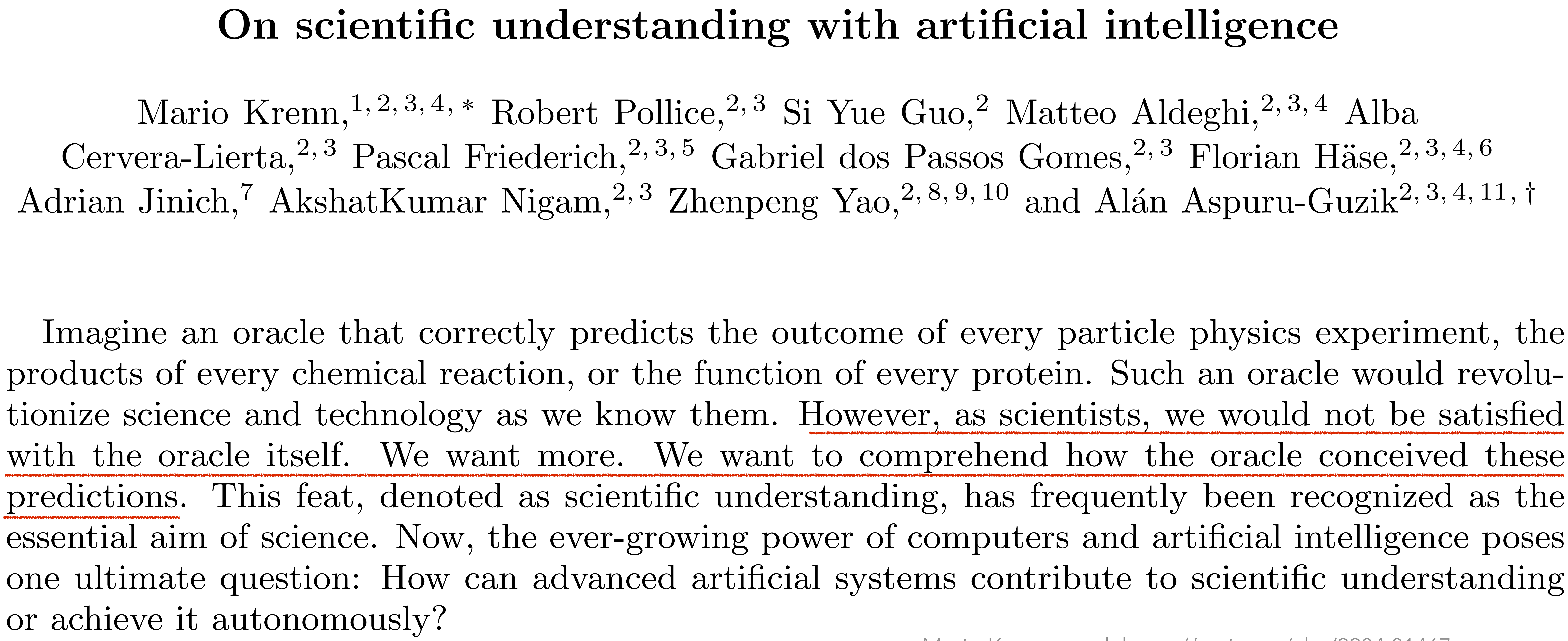
Mario Krenn,[1,2,3,4,*] Robert Pollice,[2,3] Si Yue Guo,[2] Matteo Aldeghi,[2,3,4] Alba Cervera-Lierta,[2,3] Pascal Friederich,[2,3,5] Gabriel dos Passos Gomes,[2,3] Florian Häse,[2,3,4,6] Adrian Jinich,[7] AkshatKumar Nigam,[2,3] Zhenpeng Yao,[2,8,9,10] and Alán Aspuru-Guzik[2,3,4,11,†]

Imagine an oracle that correctly predicts the outcome of every particle physics experiment, the products of every chemical reaction, or the function of every protein. Such an oracle would revolutionize science and technology as we know them. However, as scientists, we would not be satisfied with the oracle itself. We want more. We want to comprehend how the oracle conceived these predictions. This feat, denoted as scientific understanding, has frequently been recognized as the essential aim of science. Now, the ever-growing power of computers and artificial intelligence poses one ultimate question: How can advanced artificial systems contribute to scientific understanding or achieve it autonomously?

Mario Krenn, et. al. https://arxiv.org/abs/2204.01467

Scientific understanding

How does AI enable or enhance scientific understanding?



On scientific understanding with artificial intelligence

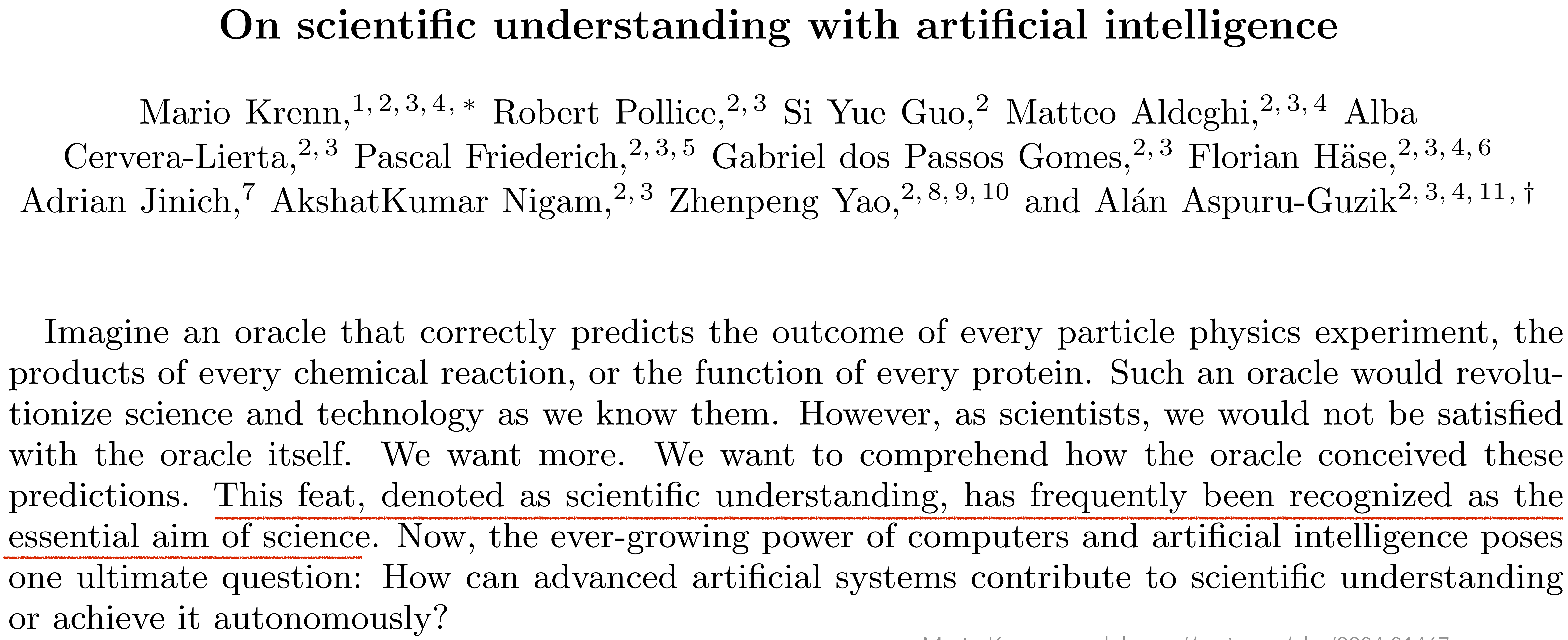
Mario Krenn,[1,2,3,4,*] Robert Pollice,[2,3] Si Yue Guo,[2] Matteo Aldeghi,[2,3,4] Alba Cervera-Lierta,[2,3] Pascal Friederich,[2,3,5] Gabriel dos Passos Gomes,[2,3] Florian Häse,[2,3,4,6] Adrian Jinich,[7] AkshatKumar Nigam,[2,3] Zhenpeng Yao,[2,8,9,10] and Alán Aspuru-Guzik[2,3,4,11,†]

Imagine an oracle that correctly predicts the outcome of every particle physics experiment, the products of every chemical reaction, or the function of every protein. Such an oracle would revolutionize science and technology as we know them. However, as scientists, we would not be satisfied with the oracle itself. We want more. We want to comprehend how the oracle conceived these predictions. This feat, denoted as scientific understanding, has frequently been recognized as the essential aim of science. Now, the ever-growing power of computers and artificial intelligence poses one ultimate question: How can advanced artificial systems contribute to scientific understanding or achieve it autonomously?

Mario Krenn, et. al. https://arxiv.org/abs/2204.01467

Scientific understanding

How does AI enable or enhance scientific understanding?



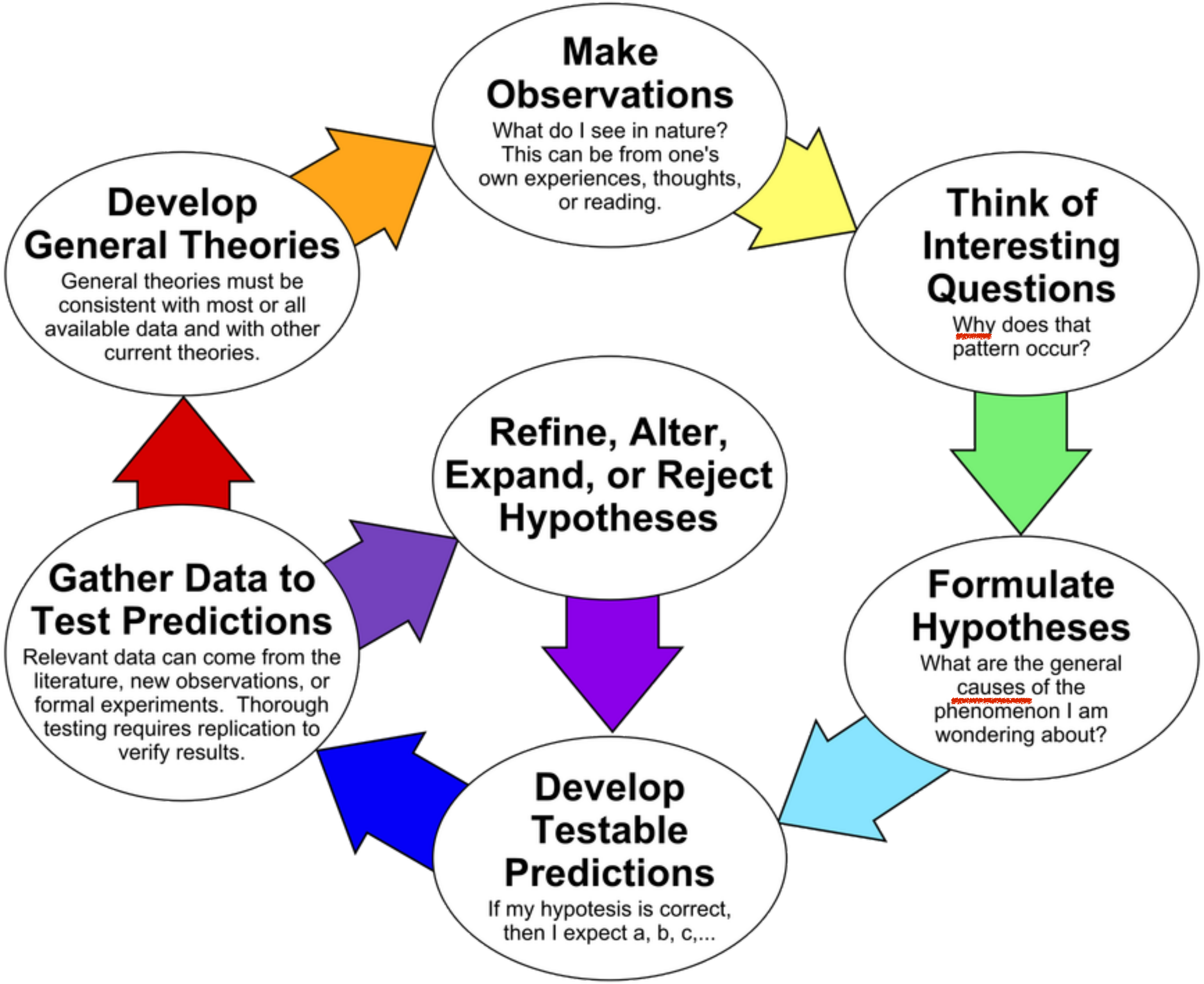
On scientific understanding with artificial intelligence

Mario Krenn,[1,2,3,4,*] Robert Pollice,[2,3] Si Yue Guo,[2] Matteo Aldeghi,[2,3,4] Alba
Cervera-Lierta,[2,3] Pascal Friederich,[2,3,5] Gabriel dos Passos Gomes,[2,3] Florian Häse,[2,3,4,6]
Adrian Jinich,[7] AkshatKumar Nigam,[2,3] Zhenpeng Yao,[2,8,9,10] and Alán Aspuru-Guzik[2,3,4,11,†]

Imagine an oracle that correctly predicts the outcome of every particle physics experiment, the products of every chemical reaction, or the function of every protein. Such an oracle would revolutionize science and technology as we know them. However, as scientists, we would not be satisfied with the oracle itself. We want more. We want to comprehend how the oracle conceived these predictions. This feat, denoted as scientific understanding, has frequently been recognized as the essential aim of science. Now, the ever-growing power of computers and artificial intelligence poses one ultimate question: How can advanced artificial systems contribute to scientific understanding or achieve it autonomously?

Mario Krenn, et. al. https://arxiv.org/abs/2204.01467

The Scientific Method as an Ongoing Process



Make Observations

What do I see in nature? This can be from one's own experiences, thoughts, or reading.

Think of Interesting Questions

Why does that pattern occur?

Formulate Hypotheses

What are the general causes of the phenomenon I am wondering about?

Develop Testable Predictions

If my hypotesis is correct, then I expect a, b, c,...

Gather Data to Test Predictions

Relevant data can come from the literature, new observations, or formal experiments. Thorough testing requires replication to verify results.

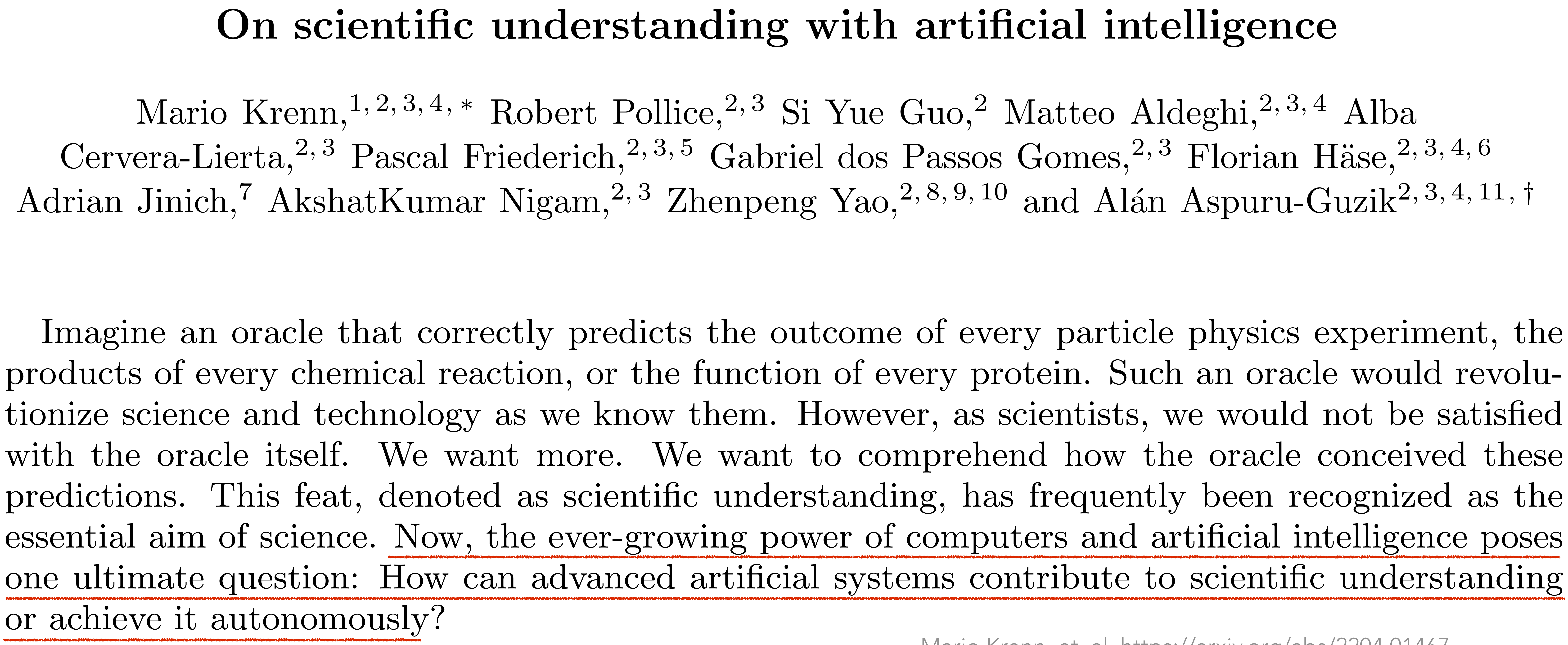
Refine, Alter, Expand, or Reject Hypotheses

Develop General Theories

General theories must be consistent with most or all available data and with other current theories.

Looking forward

How does AI enable or enhance scientific understanding?



On scientific understanding with artificial intelligence

Mario Krenn,[1,2,3,4,*] Robert Pollice,[2,3] Si Yue Guo,[2] Matteo Aldeghi,[2,3,4] Alba Cervera-Lierta,[2,3] Pascal Friederich,[2,3,5] Gabriel dos Passos Gomes,[2,3] Florian Häse,[2,3,4,6] Adrian Jinich,[7] AkshatKumar Nigam,[2,3] Zhenpeng Yao,[2,8,9,10] and Alán Aspuru-Guzik[2,3,4,11,†]

Imagine an oracle that correctly predicts the outcome of every particle physics experiment, the products of every chemical reaction, or the function of every protein. Such an oracle would revolutionize science and technology as we know them. However, as scientists, we would not be satisfied with the oracle itself. We want more. We want to comprehend how the oracle conceived these predictions. This feat, denoted as scientific understanding, has frequently been recognized as the essential aim of science. Now, the ever-growing power of computers and artificial intelligence poses one ultimate question: How can advanced artificial systems contribute to scientific understanding or achieve it autonomously?

Mario Krenn, et. al. https://arxiv.org/abs/2204.01467

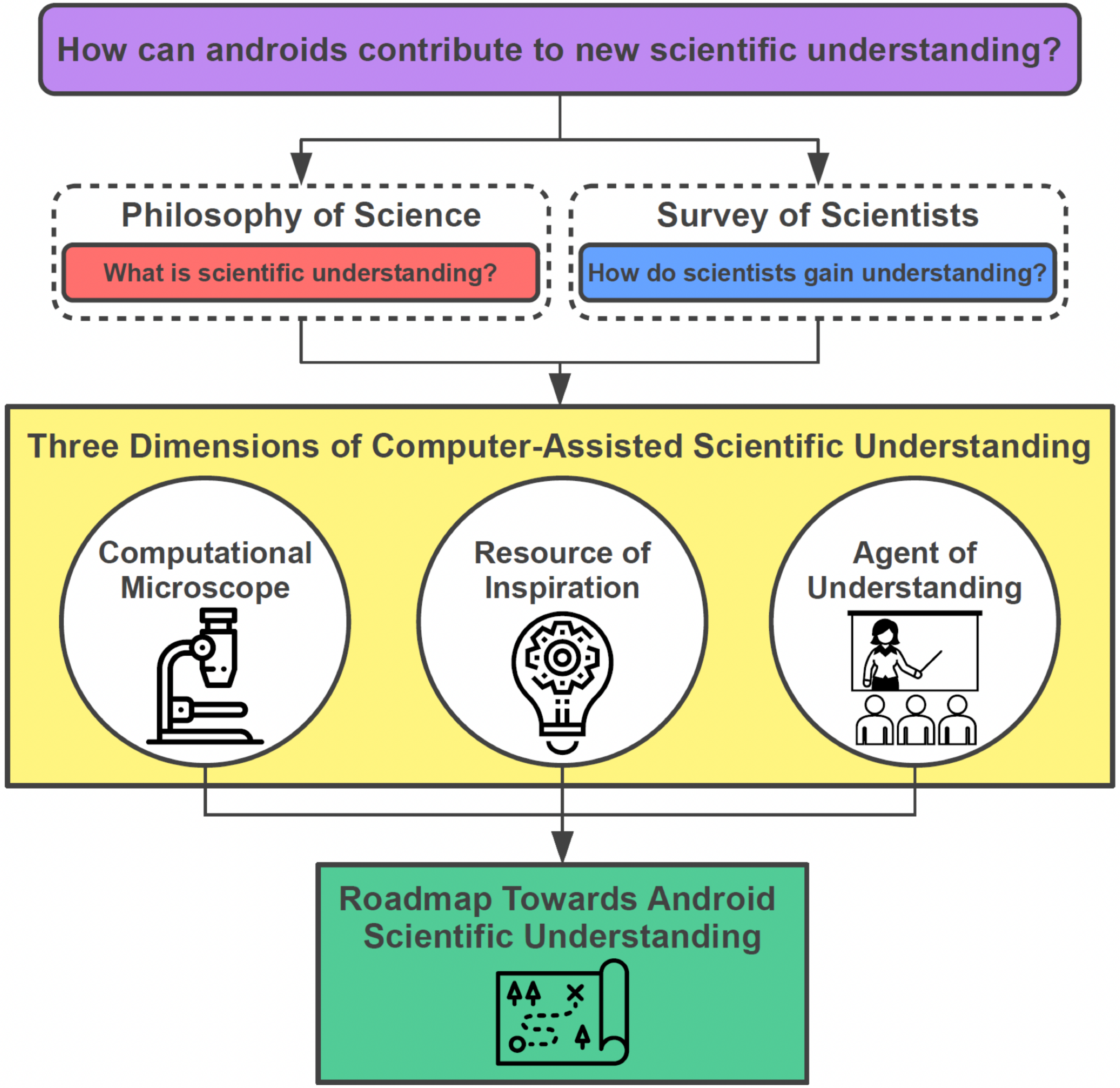
How does AI help?

An android can act

- I) as a **computational microscope**, providing information not (yet) attainable by experiment
- II) as a **resource of inspiration** or an artificial muse, expanding the scope of human imagination and creativity.

In those two classes, the human scientist is essential to take the new insight and inspiration and develop it to full understanding. Finally, an android can be

- III) an **agent of understanding**, replacing the human in generalizing observations and transferring scientific concepts to new phenomena.



How can androids contribute to new scientific understanding?

Philosophy of Science

What is scientific understanding?

Survey of Scientists

How do scientists gain understanding?

Three Dimensions of Computer-Assisted Scientific Understanding

Computational Microscope

Resource of Inspiration

Agent of Understanding

Roadmap Towards Android Scientific Understanding

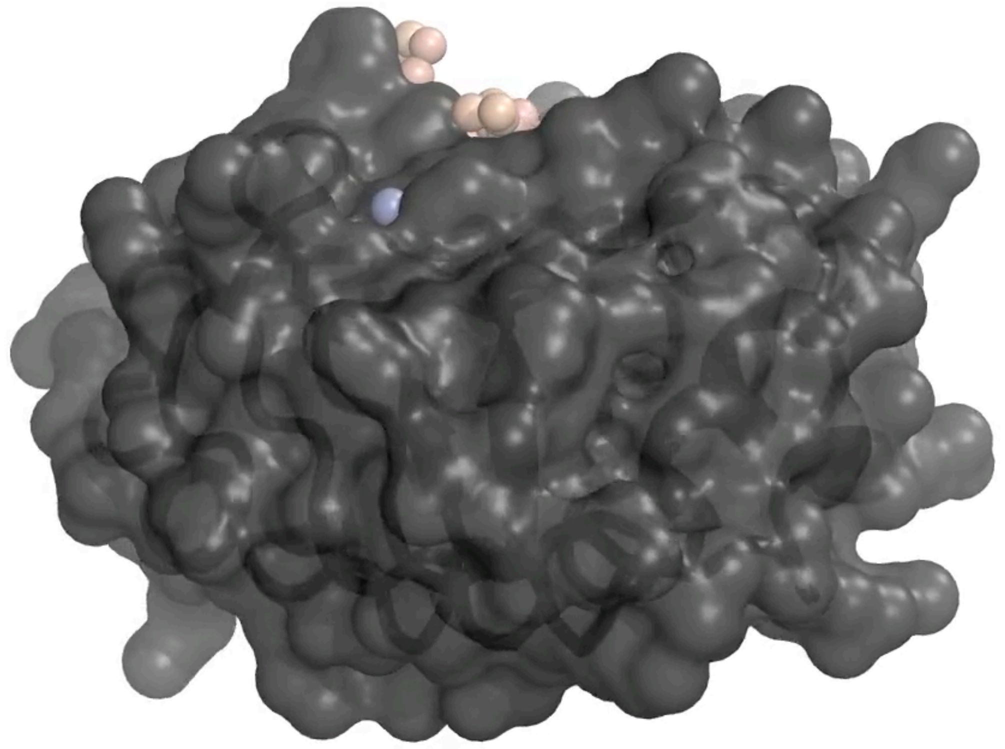
Mario Krenn, et. Al. https://arxiv.org/abs/2204.01467

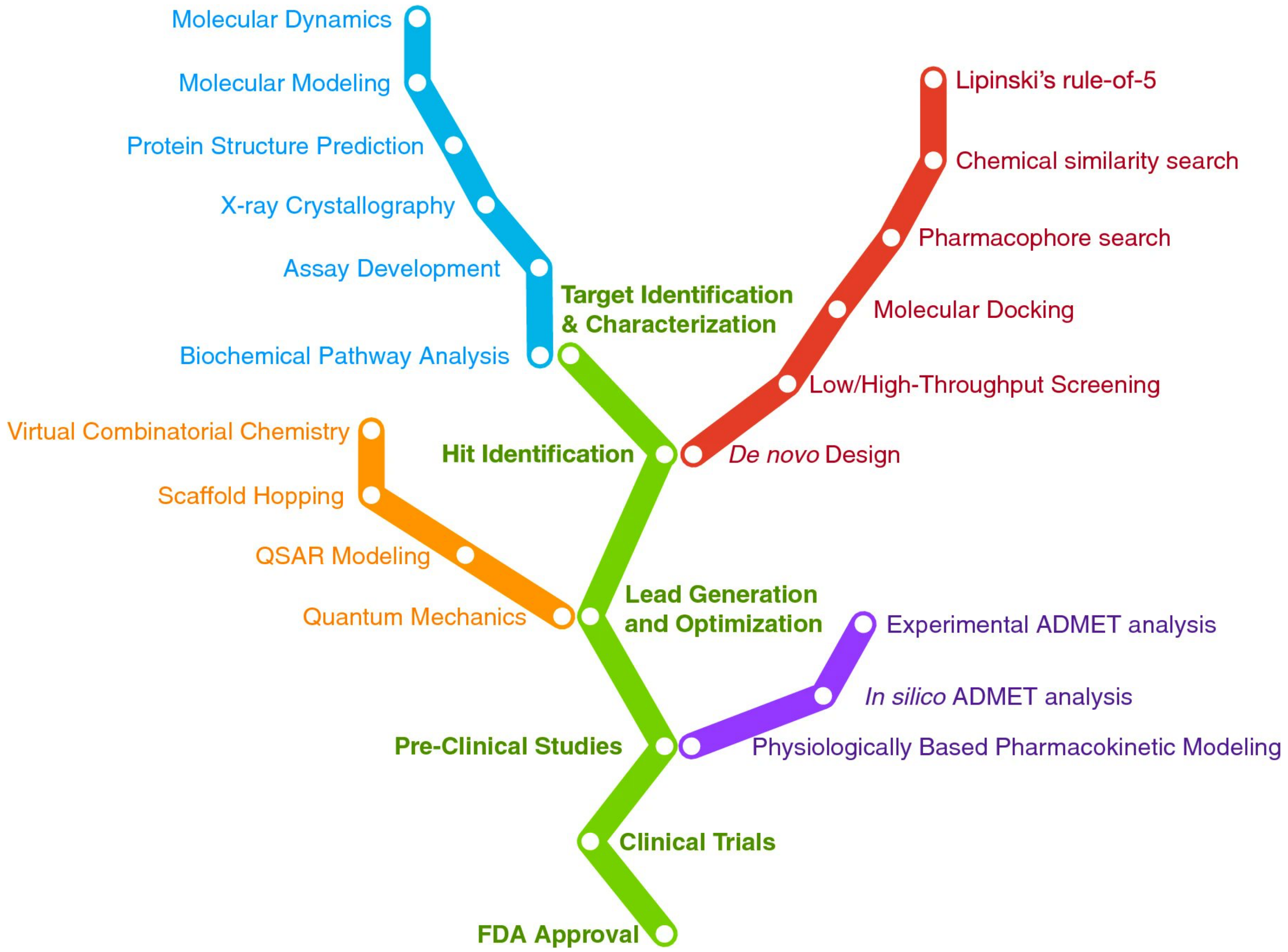
Use of ML in
Physics vs. Molecules & Materials

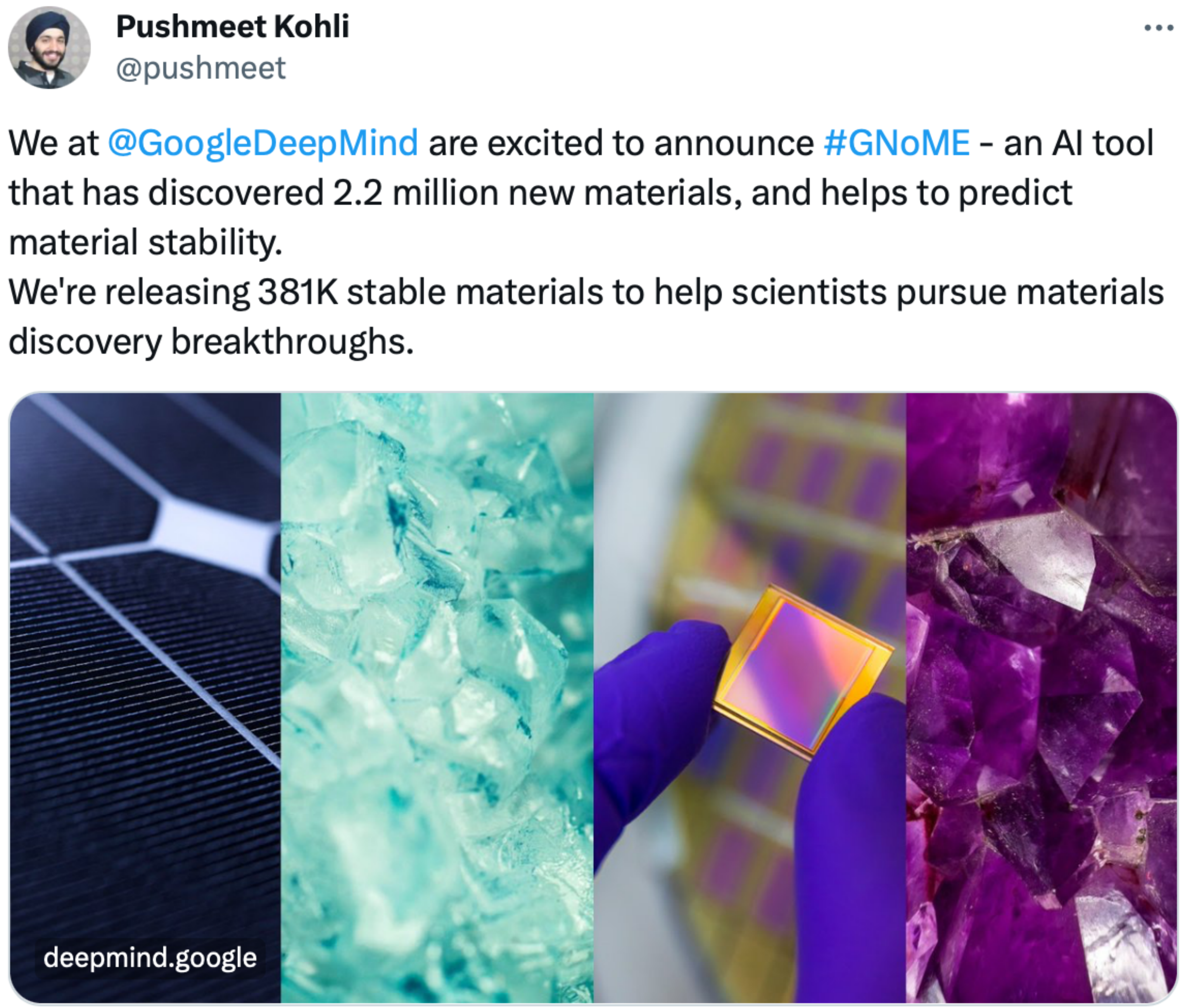
Drug & Materials Discovery

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







Pushmeet Kohli
@pushmeet

We at @GoogleDeepMind are excited to announce #GNoME - an AI tool that has discovered 2.2 million new materials, and helps to predict material stability.
We're releasing 381K stable materials to help scientists pursue materials discovery breakthroughs.

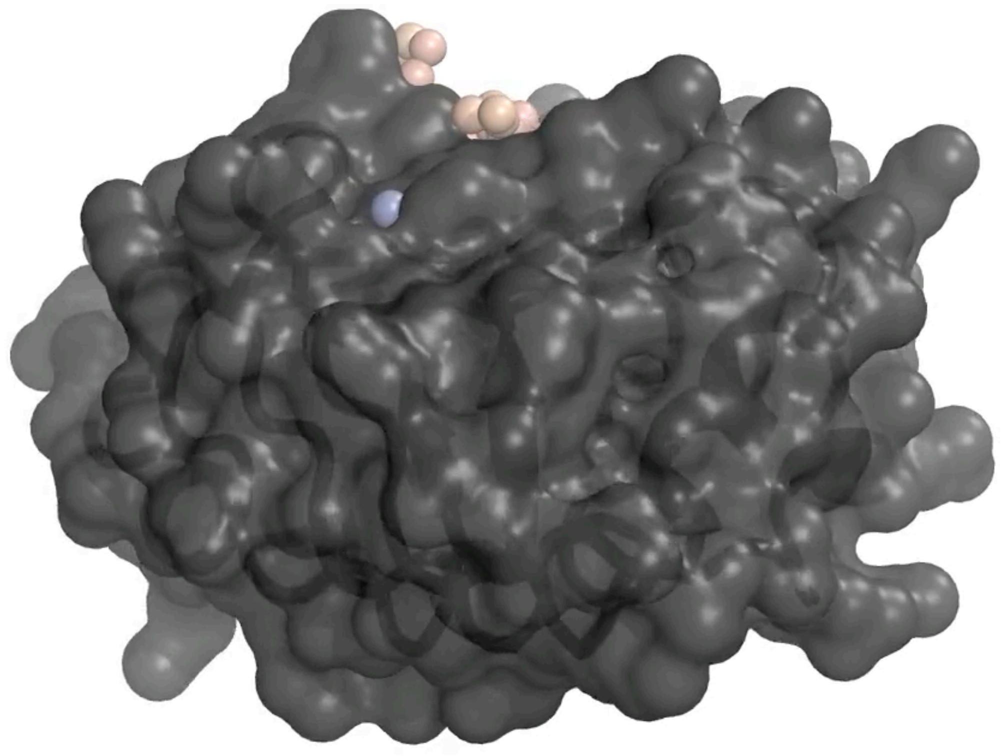
Image credits: RFDiffusion from Baker Lab, Institute for Protein Design, U Washington;
Chanin Nantasenamat from Towards reproducible computational drug discovery. *J Cheminform* **12**, 9 (2020). https://doi.org/10.1186/s13321-020-0408-x

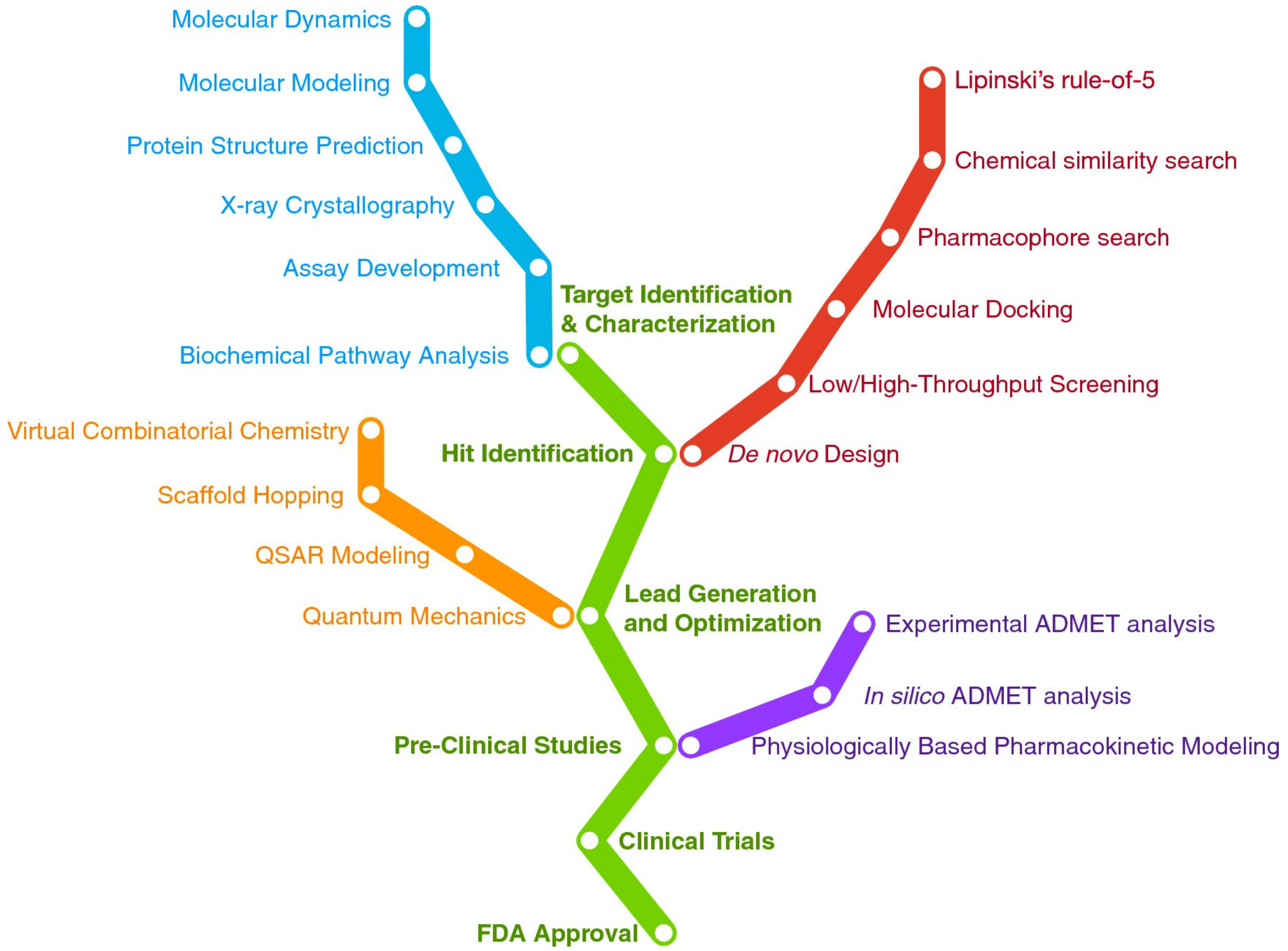
https://dpmd.ai/PK-materials

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





Molecular Dynamics
Molecular Modeling
Protein Structure Prediction
X-ray Crystallography
Assay Development
Biochemical Pathway Analysis

Target Identification & Characterization

Lipinski's rule-of-5
Chemical similarity search
Pharmacophore search
Molecular Docking
Low/High-Throughput Screening
De novo Design

Hit Identification

Virtual Combinatorial Chemistry
Scaffold Hopping
QSAR Modeling
Quantum Mechanics

Lead Generation and Optimization

Experimental ADMET analysis
In silico ADMET analysis
Physiologically Based Pharmacokinetic Modeling

Pre-Clinical Studies

Clinical Trials

FDA Approval

Pushmeet Kohli
@pushmeet

We at @GoogleDeepMind are excited to announce #GNoME - an AI tool that has discovered 2.2 million new materials, and helps to predict material stability.
We're releasing 381K stable materials to help scientists pursue materials discovery breakthroughs.

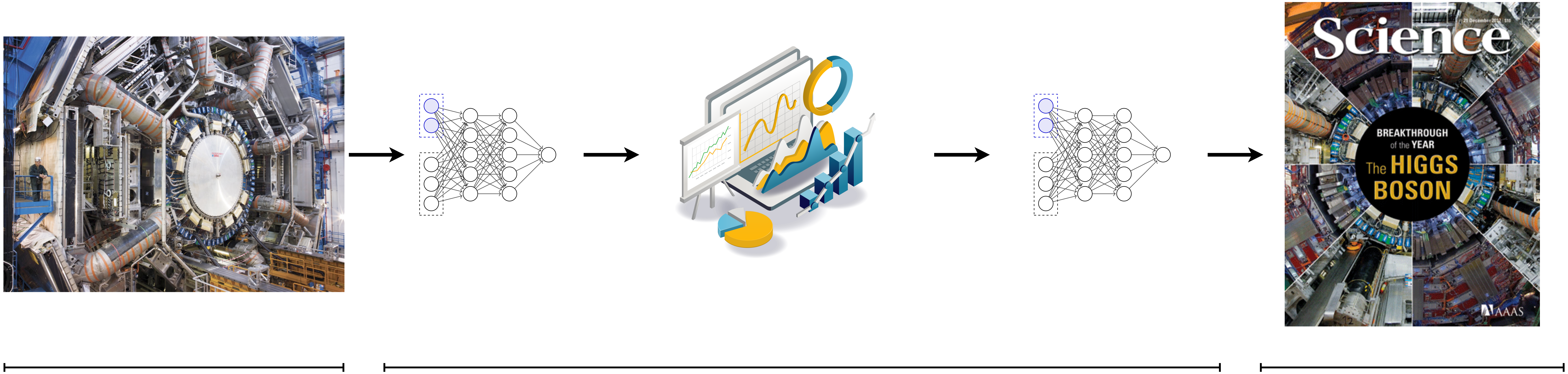
Image credits: RFDiffusion from Baker Lab, Institute for Protein Design, U Washington;
Chanin Nantasenamat from Towards reproducible computational drug discovery. *J Cheminform* **12**, 9 (2020). https://doi.org/10.1186/s13321-020-0408-x

https://dpmd.ai/PK-materials

Experimental Physics, Astrophysics, Cosmology

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

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

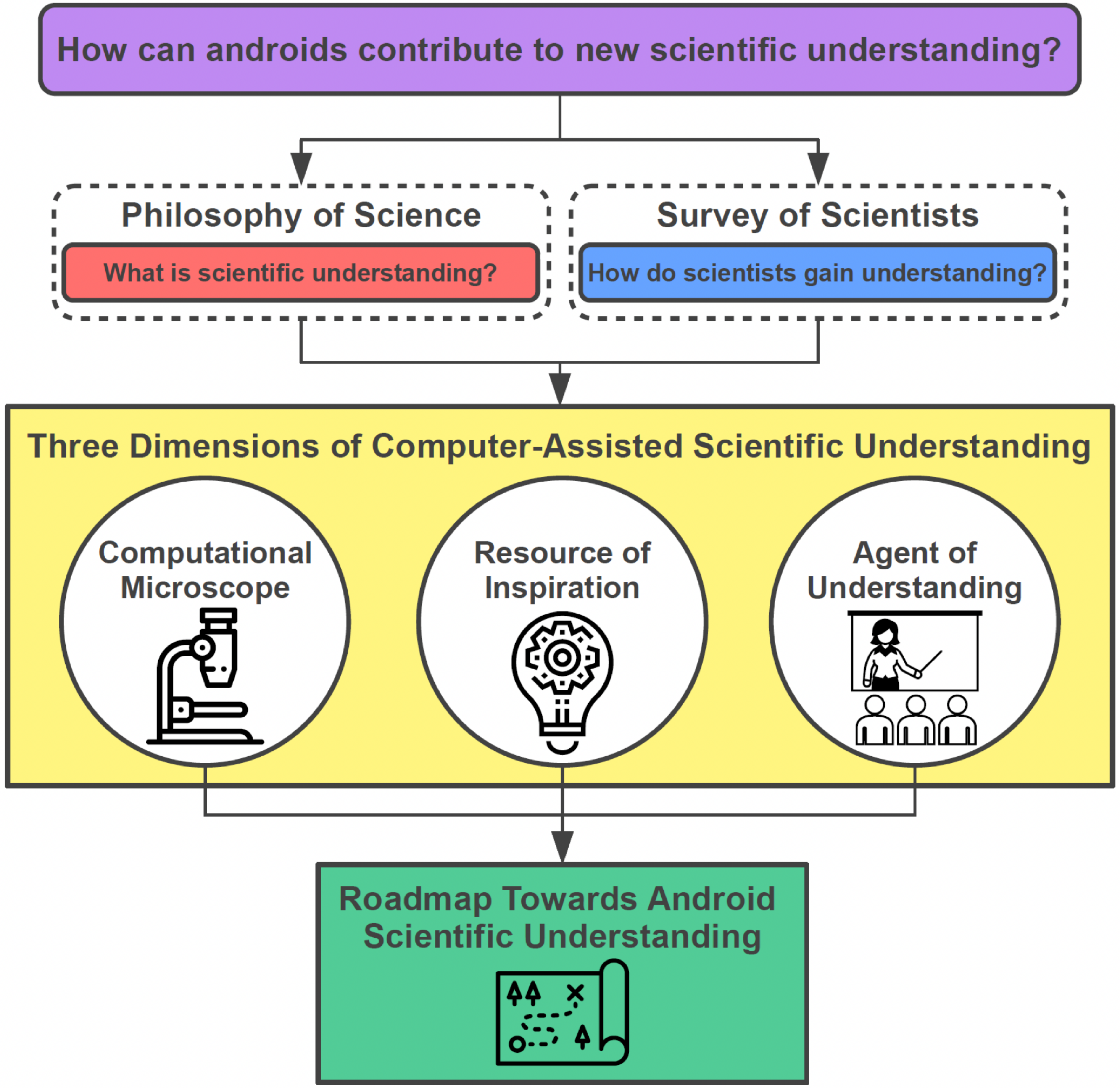


Data Collection

Data Analysis Pipeline

Scientific Claims

Computer-Assisted Scientific Understanding



How can androids contribute to new scientific understanding?

Philosophy of Science

What is scientific understanding?

Survey of Scientists

How do scientists gain understanding?

Three Dimensions of Computer-Assisted Scientific Understanding

Computational Microscope

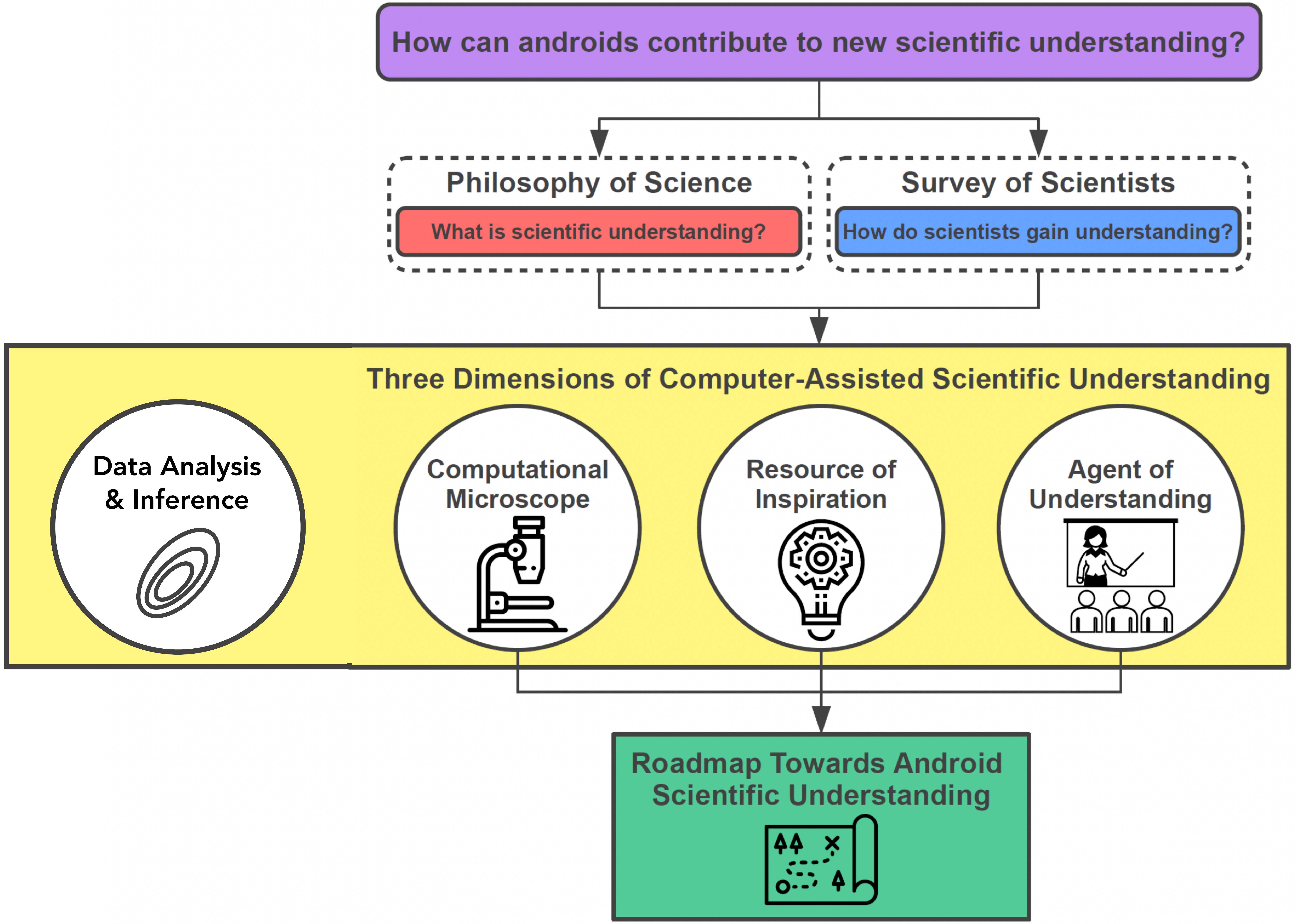
Resource of Inspiration

Agent of Understanding

Roadmap Towards Android Scientific Understanding

Mario Krenn, et. Al. https://arxiv.org/abs/2204.01467

Computer-Assisted Scientific Understanding



How can androids contribute to new scientific understanding?

Philosophy of Science

What is scientific understanding?

Survey of Scientists

How do scientists gain understanding?

Three Dimensions of Computer-Assisted Scientific Understanding

Data Analysis & Inference

Computational Microscope

Resource of Inspiration

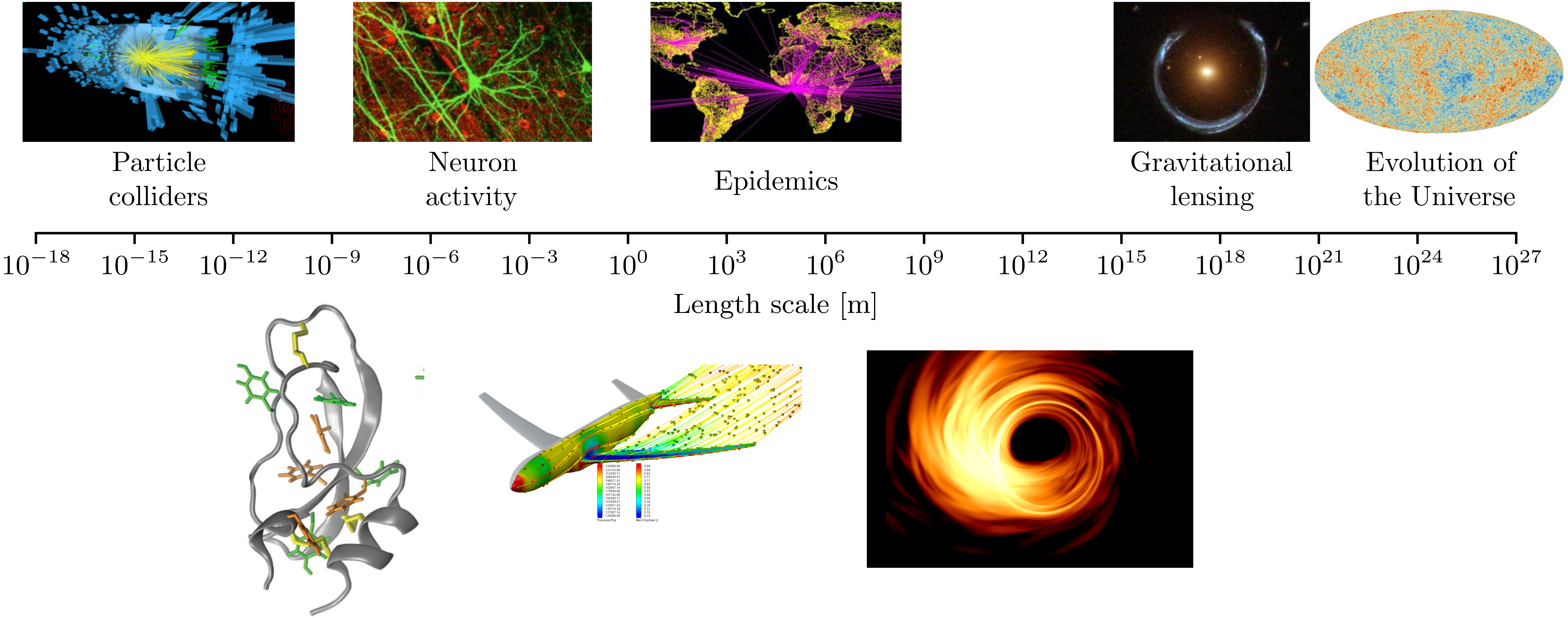
Agent of Understanding

Roadmap Towards Android Scientific Understanding

Mario Krenn, et. Al. https://arxiv.org/abs/2204.01467

Simulation-based Inference

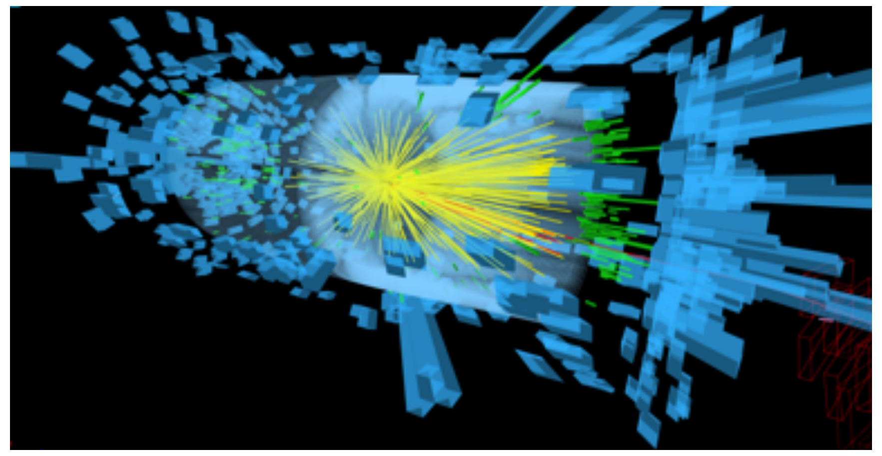
Science is replete with high-fidelity simulators

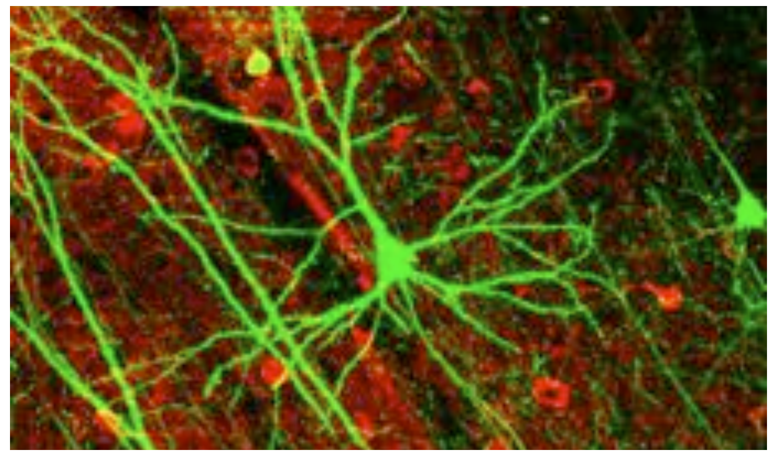


Simulators are causal, generative models of the data generating process

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

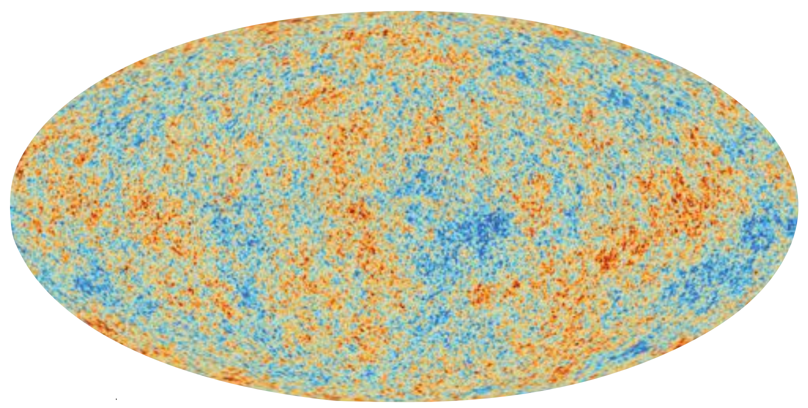
Science is replete with high-fidelity simulators

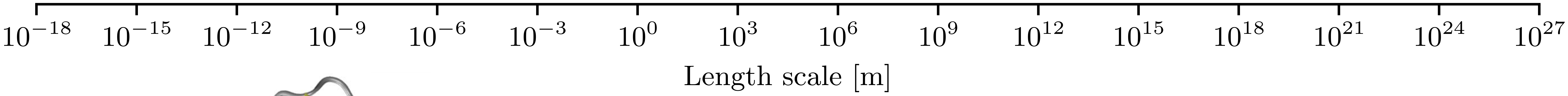

Particle colliders

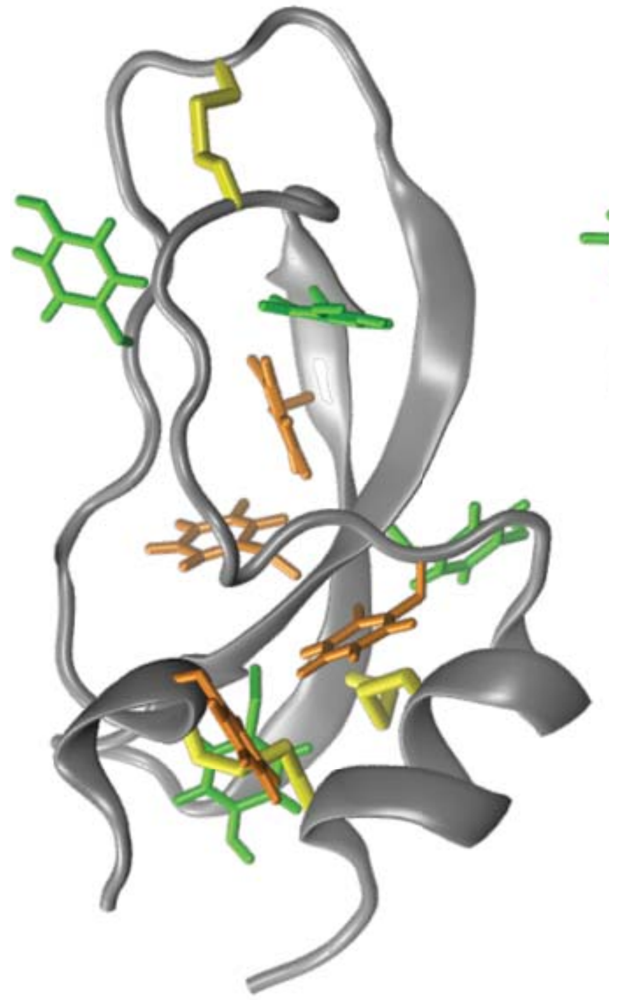

Neuron activity

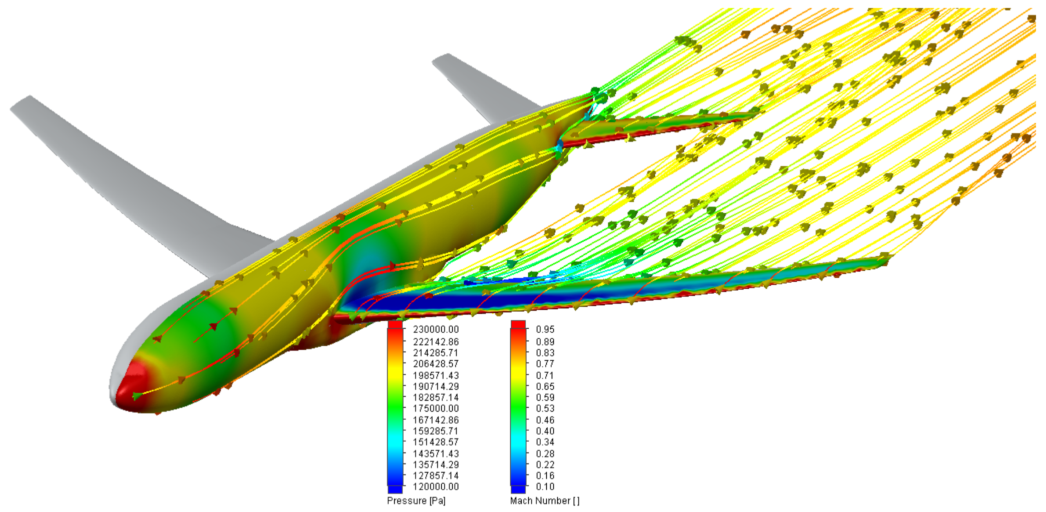

Epidemics

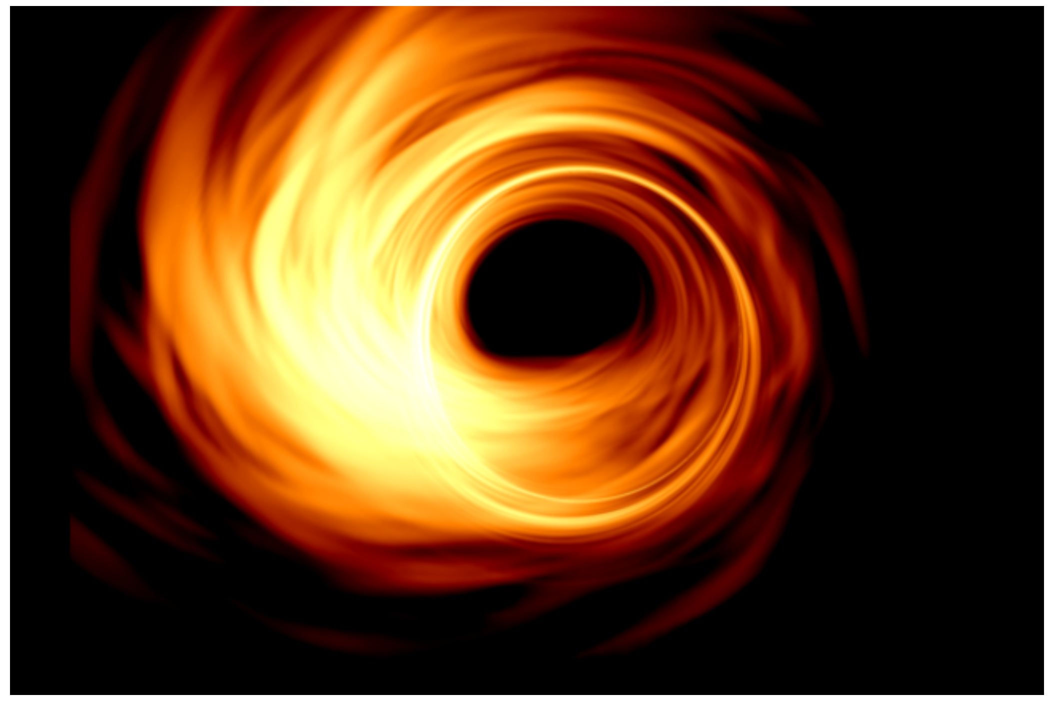

Gravitational lensing


Evolution of the Universe





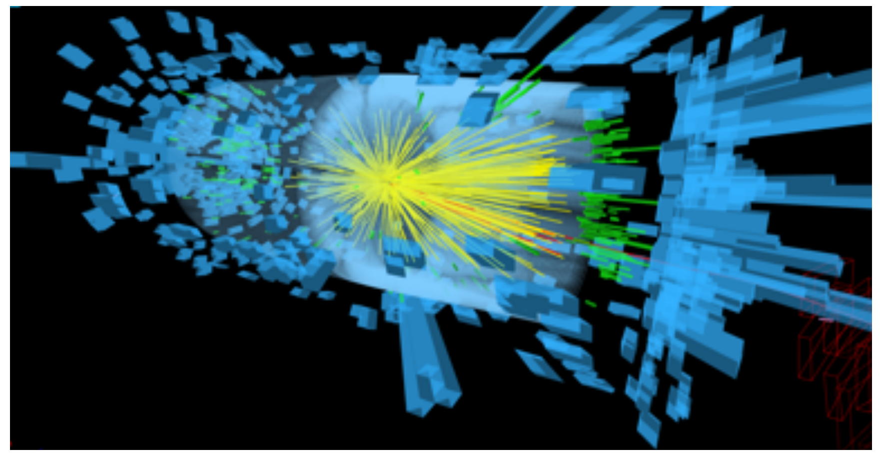


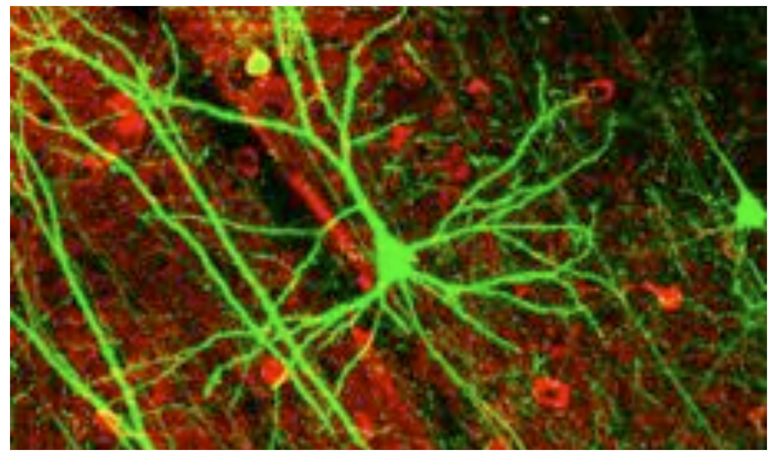


The expressiveness of programming languages facilitates the development of complex, high-fidelity simulations, and the power of modern computing provides the ability to generate synthetic data from them.

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

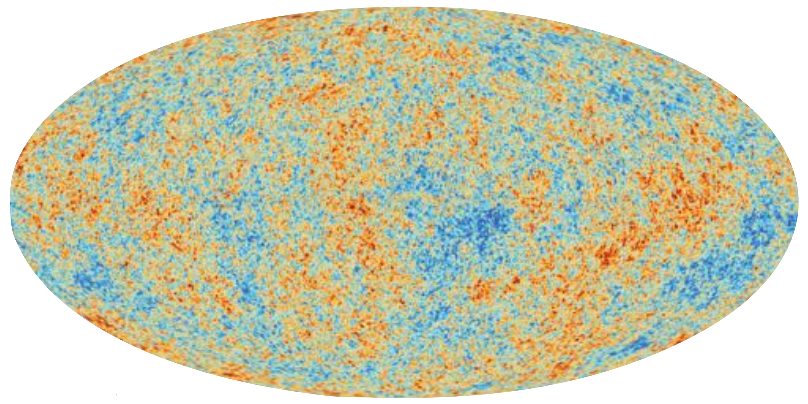
Science is replete with high-fidelity simulators

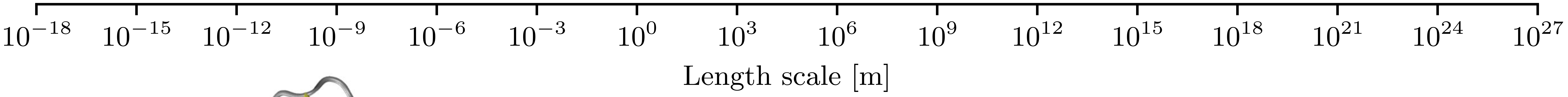

Particle colliders

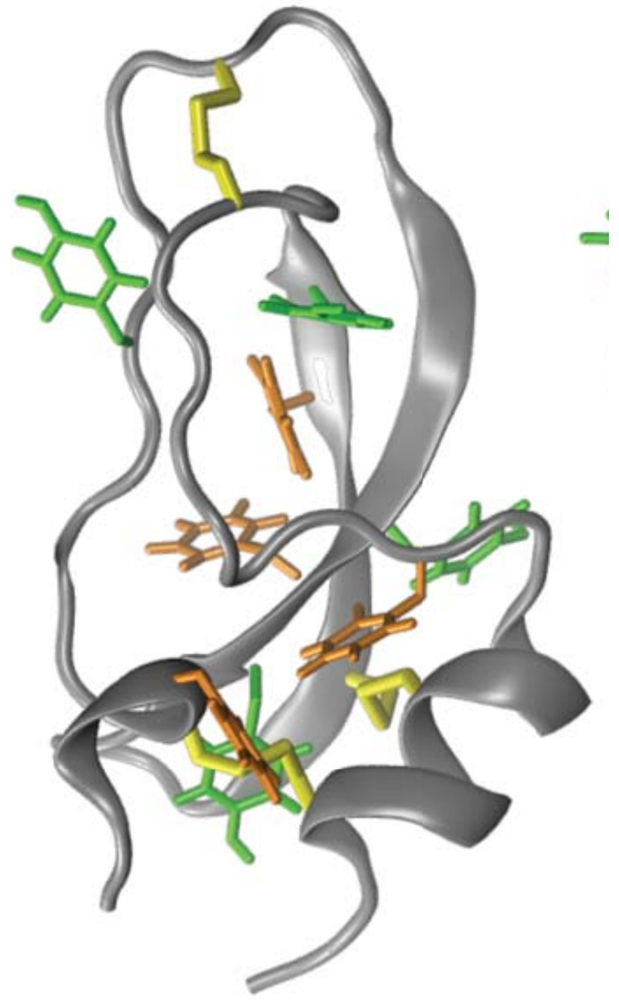

Neuron activity

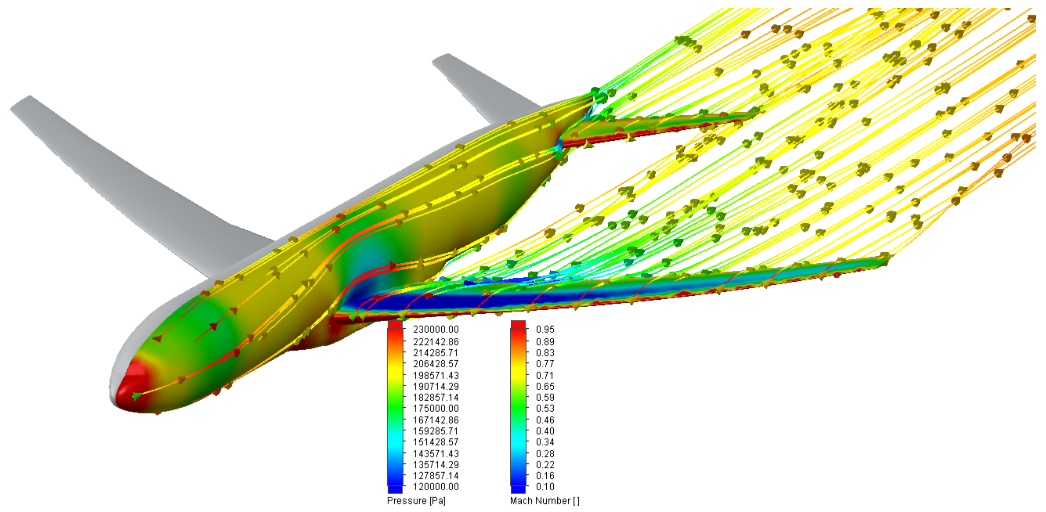

Epidemics

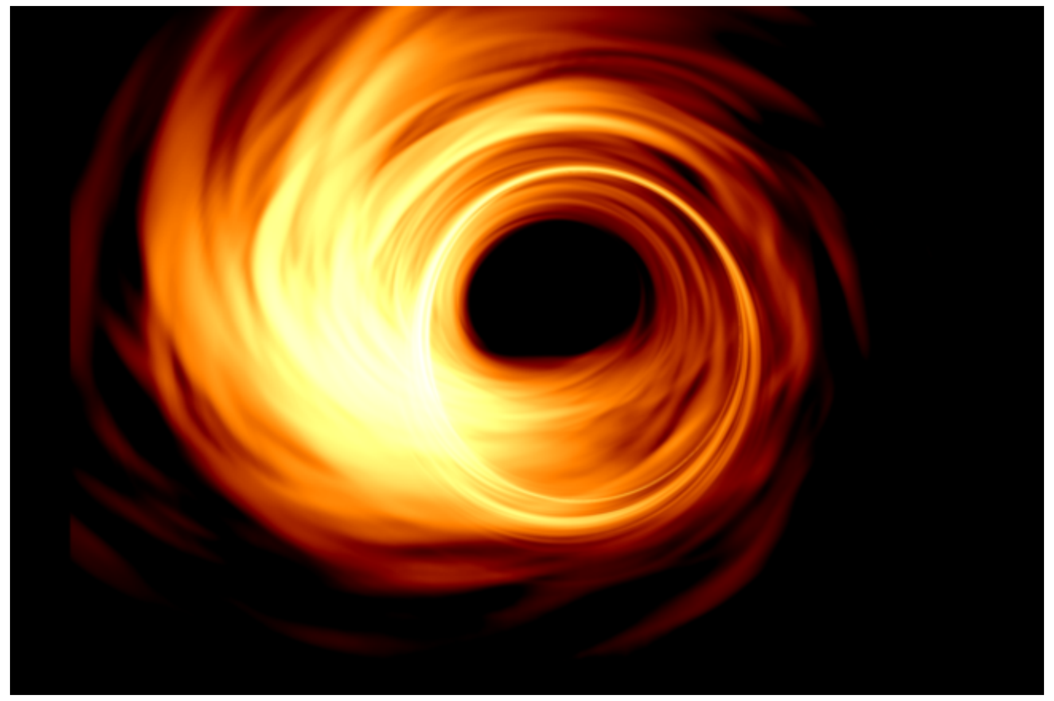

Gravitational lensing


Evolution of the Universe







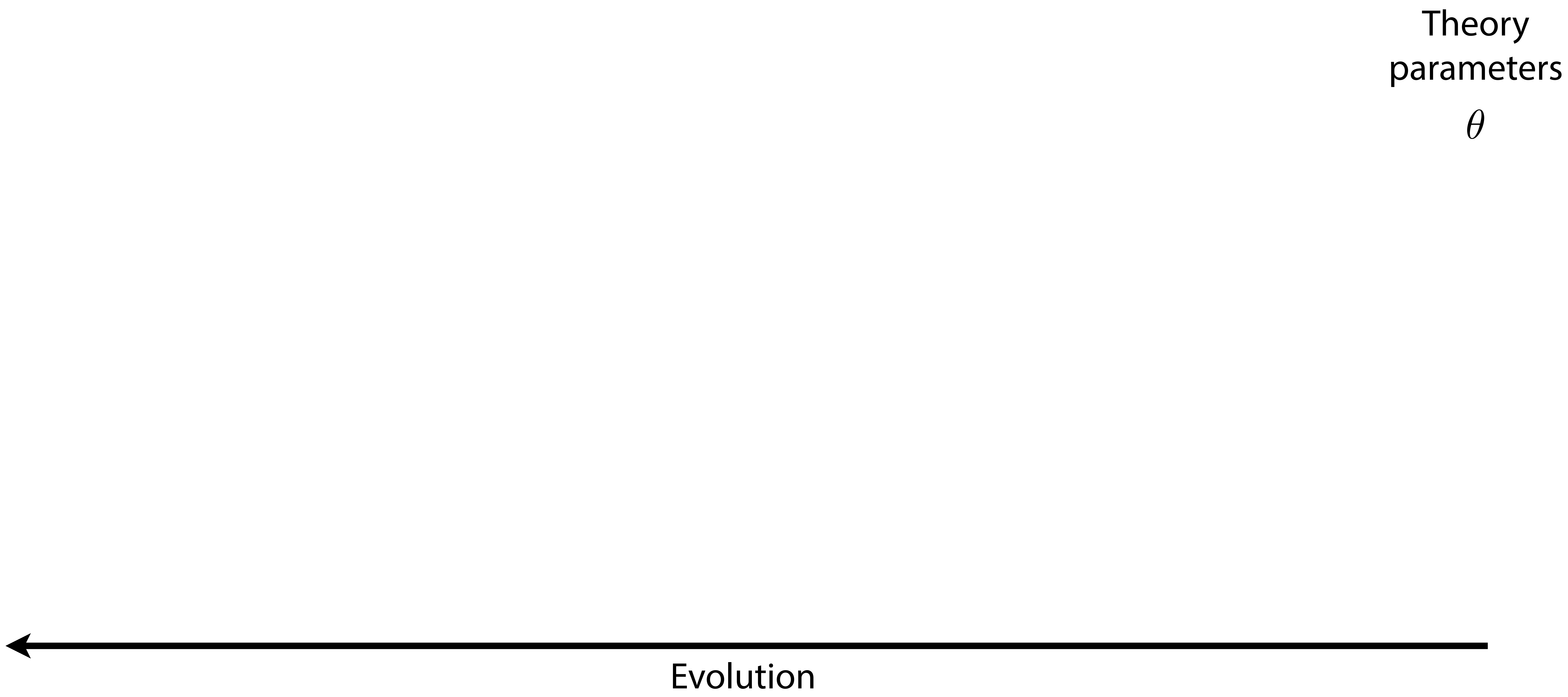


Unfortunately, these simulators are poorly suited for inference.

Often involve many latent variables, non-differentiable components, or they are just too slow.

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

Simulating particle physics processes



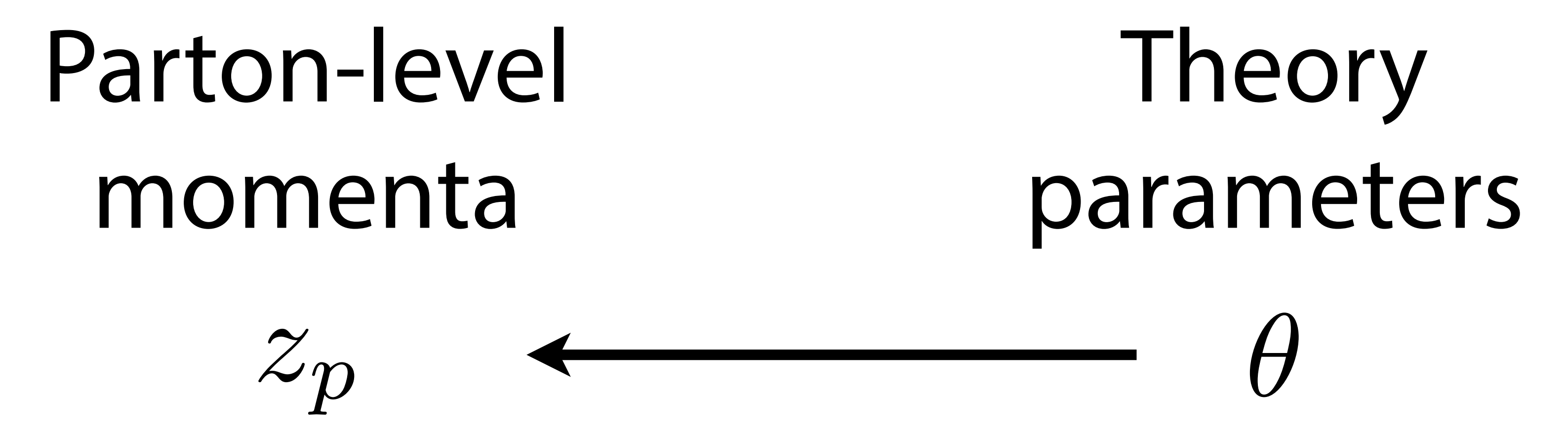
Theory parameters

θ

Evolution

Simulating particle physics processes

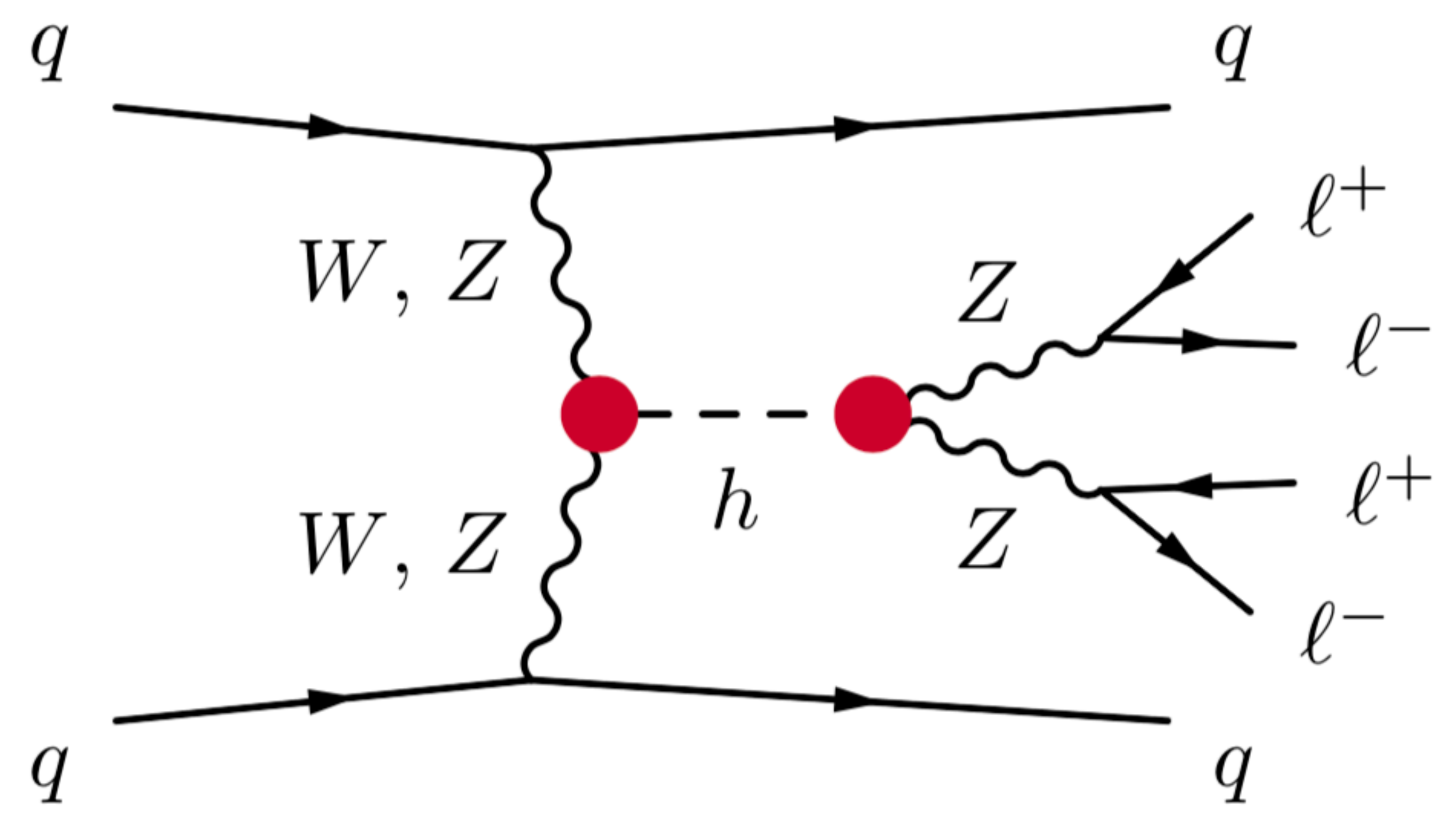
Latent variables



Parton-level momenta

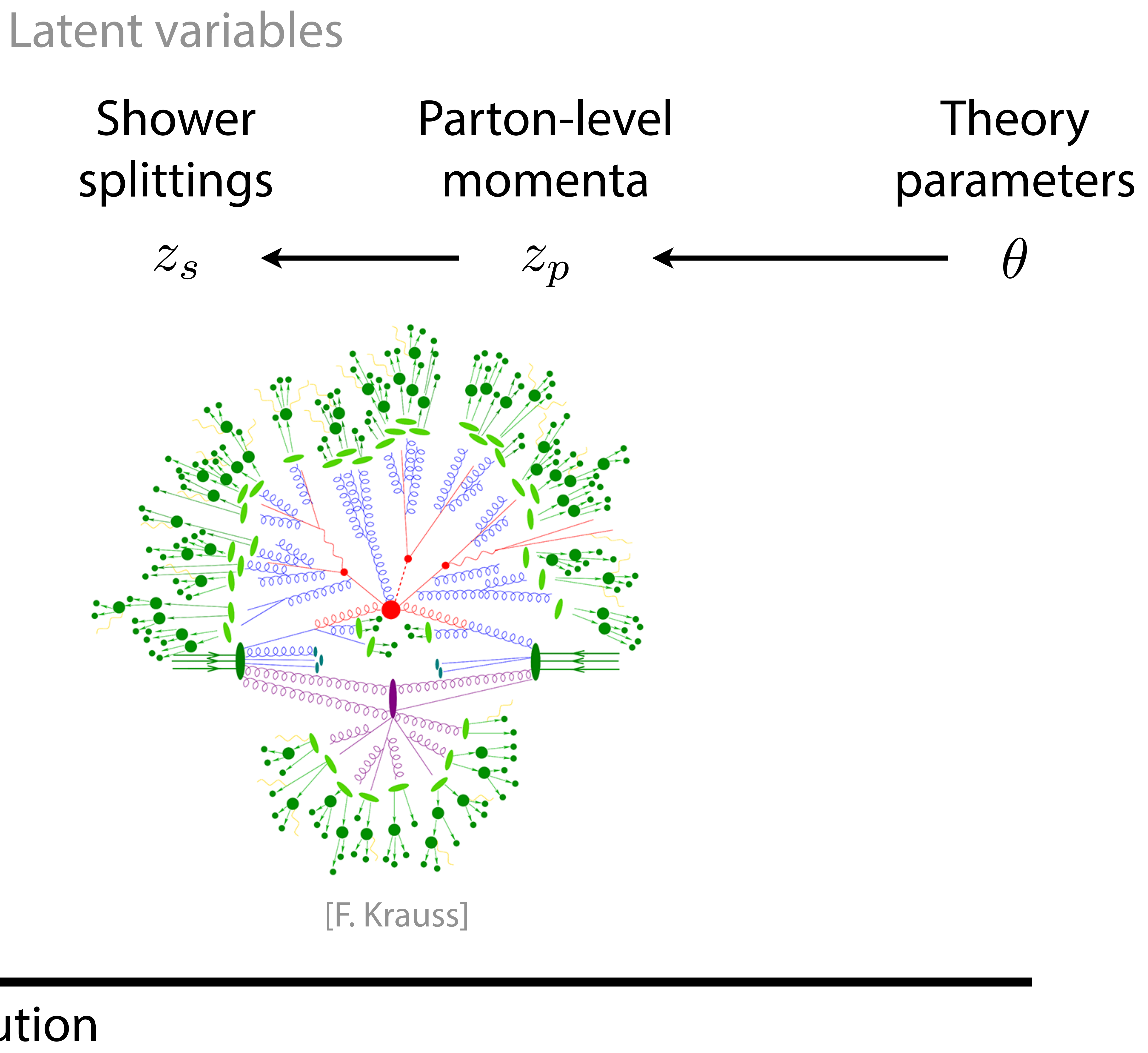
Theory parameters

$z_p \longleftarrow \theta$



Evolution

Simulating particle physics processes



Latent variables

Shower splittings

Parton-level momenta

Theory parameters

$z_s \longleftarrow z_p \longleftarrow \theta$

[F. Krauss]

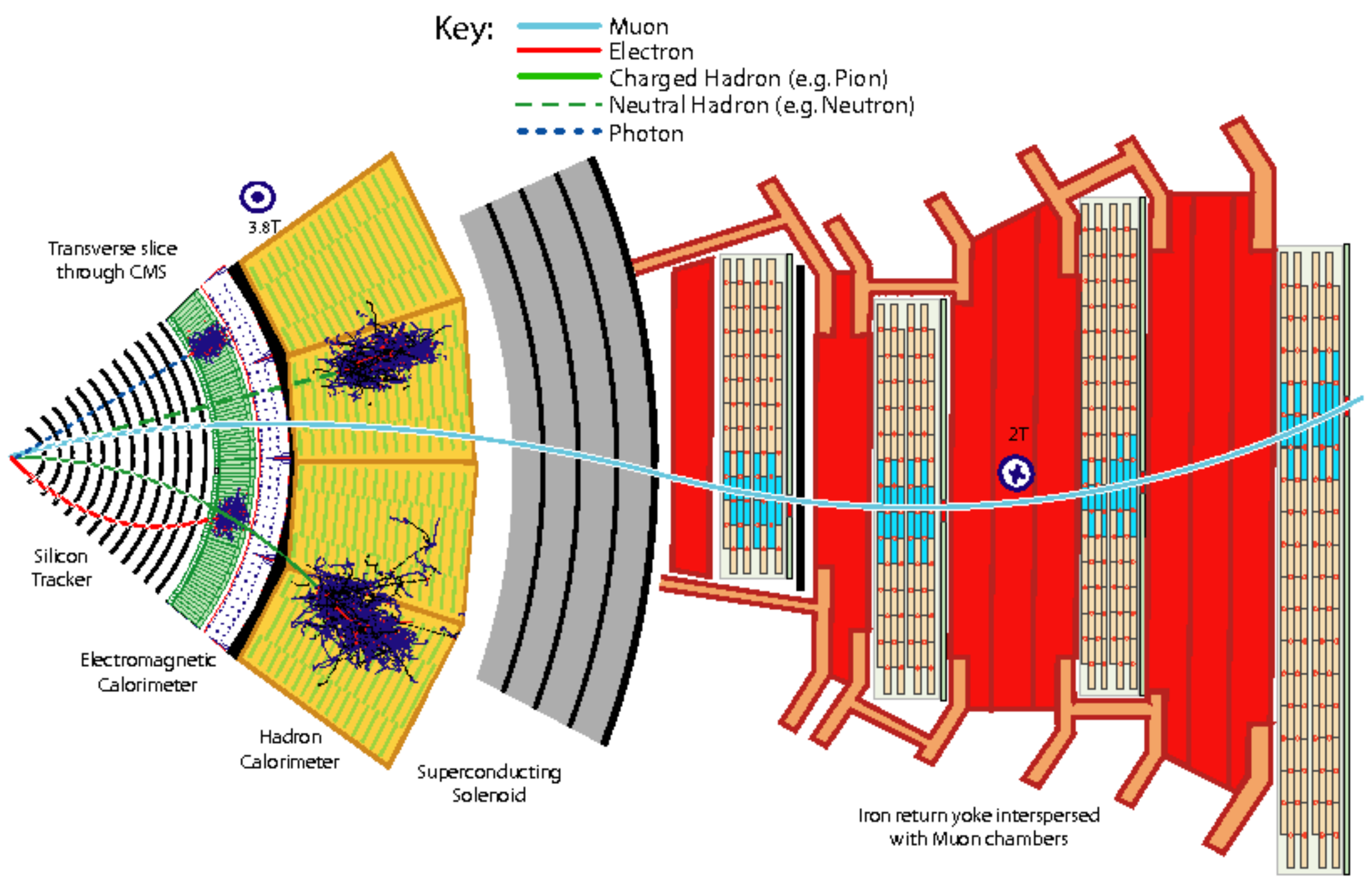
Evolution

Simulating particle physics processes

Latent variables

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

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



[CMS]

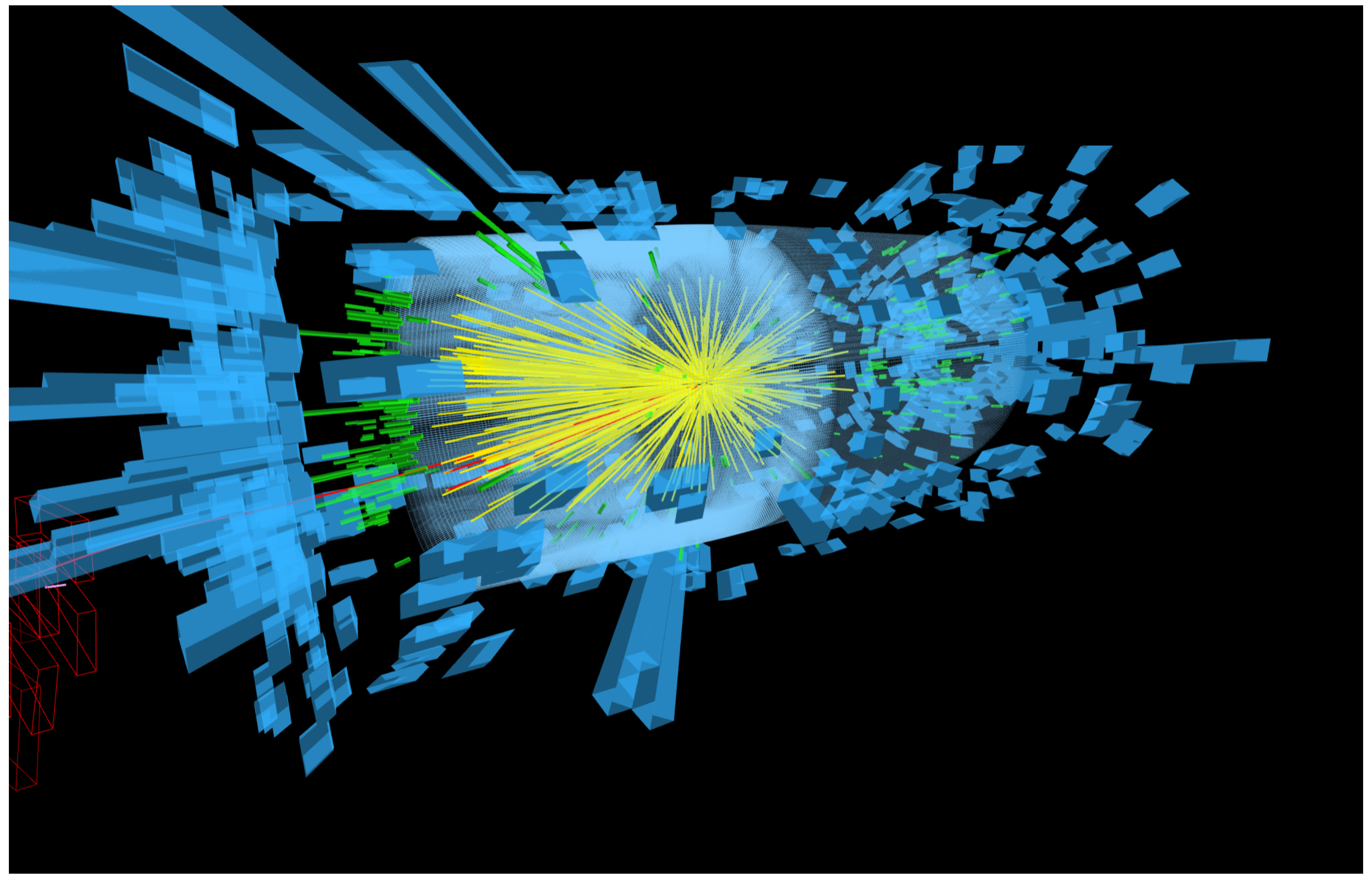
$\longleftarrow$ Evolution

Simulating particle physics processes

Latent variables

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

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



$\longleftarrow$

Evolution

Simulating particle physics processes

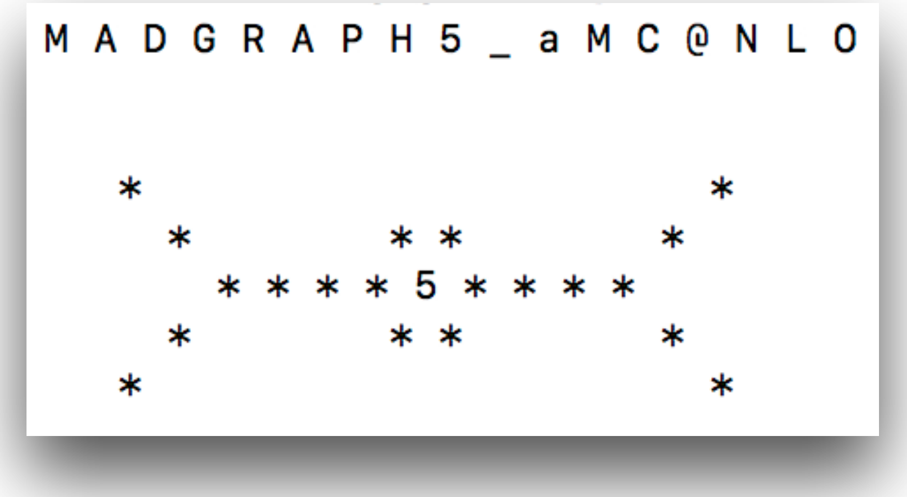
Latent variables

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

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

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

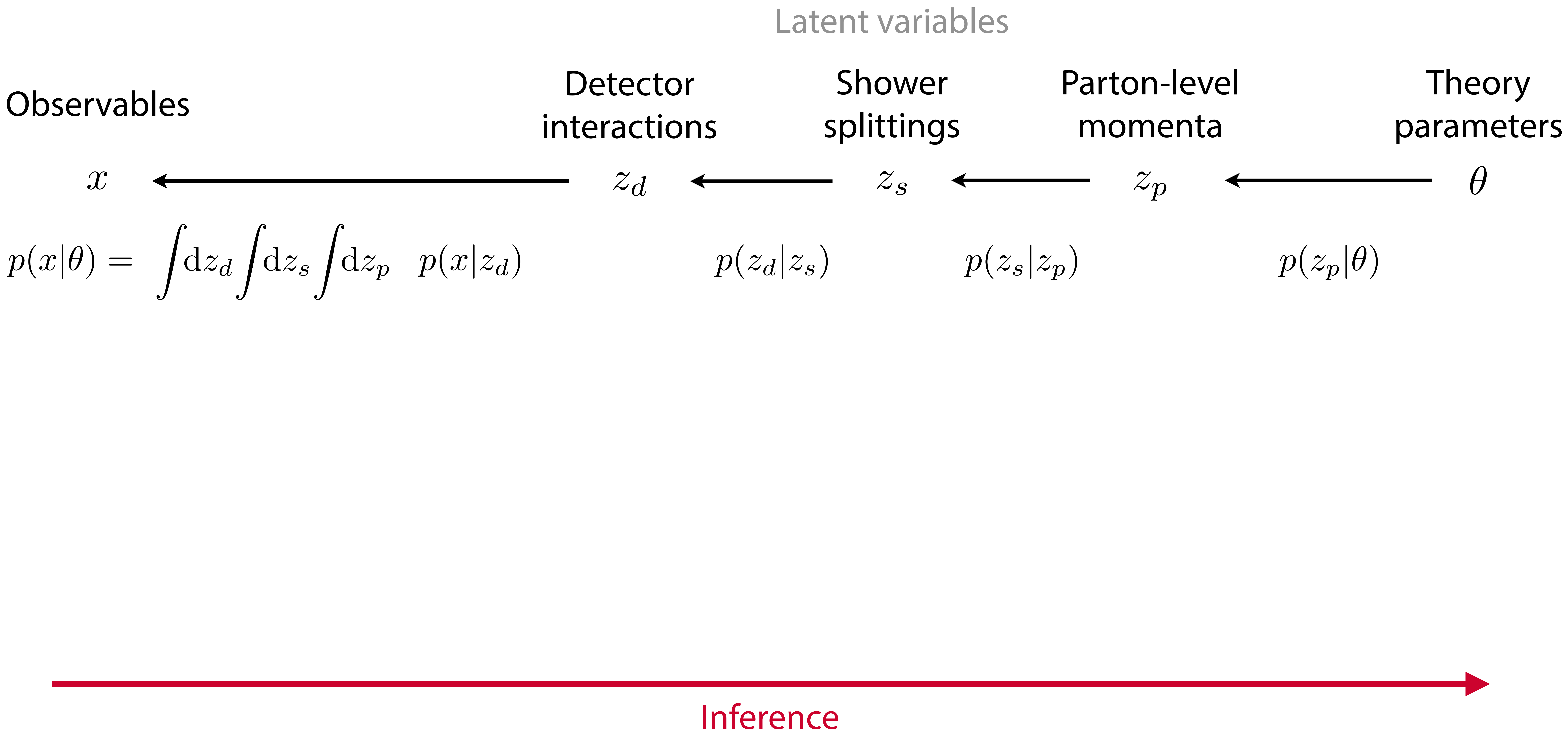
GEANT4
A SIMULATION TOOLKIT



```
M A D G R A P H 5 _ a M C @ N L O

*                     *
  *        * *      *
    * * * * 5 * * * *
  *        * *      *
*                     *
```

Prediction (simulation)

Simulating particle physics processes



Latent variables

Observables

Detector interactions

Shower splittings

Parton-level momenta

Theory parameters

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

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

Inference

Simulating particle physics processes

Latent variables

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

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

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

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

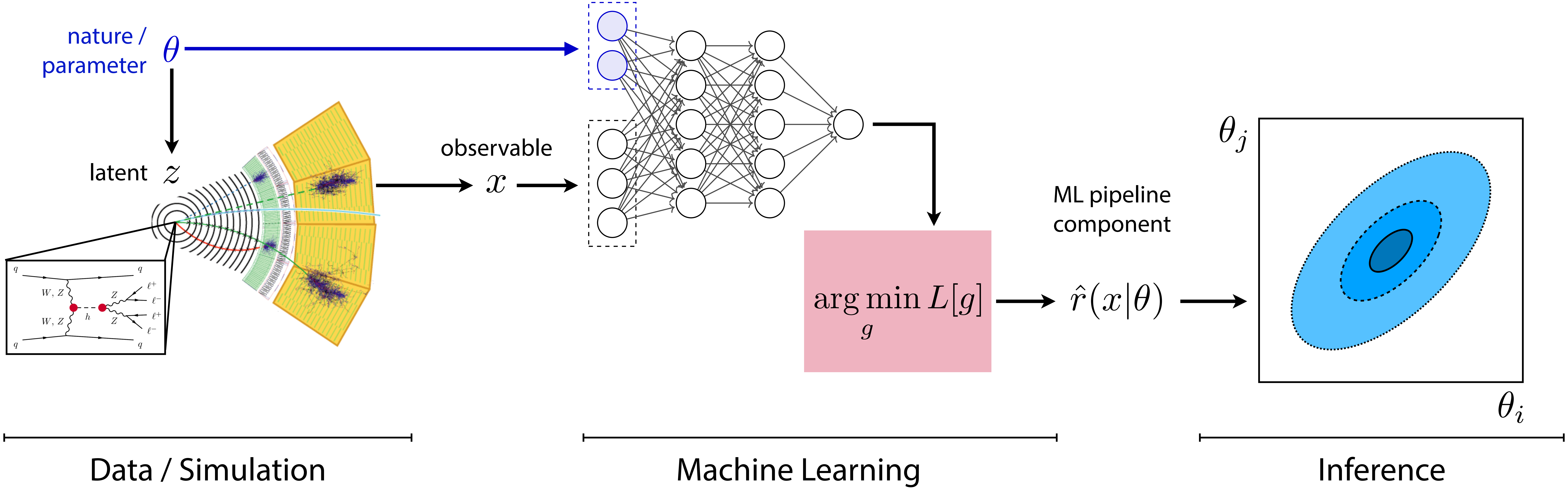
$$\xrightarrow{\hspace{10cm}}$$

Inference

Simulation-Based Inference

Deep learning and neural density estimation are effective at learning approximate surrogates for the likelihood and posterior, **revolutionizing principled statistical inference in science!**

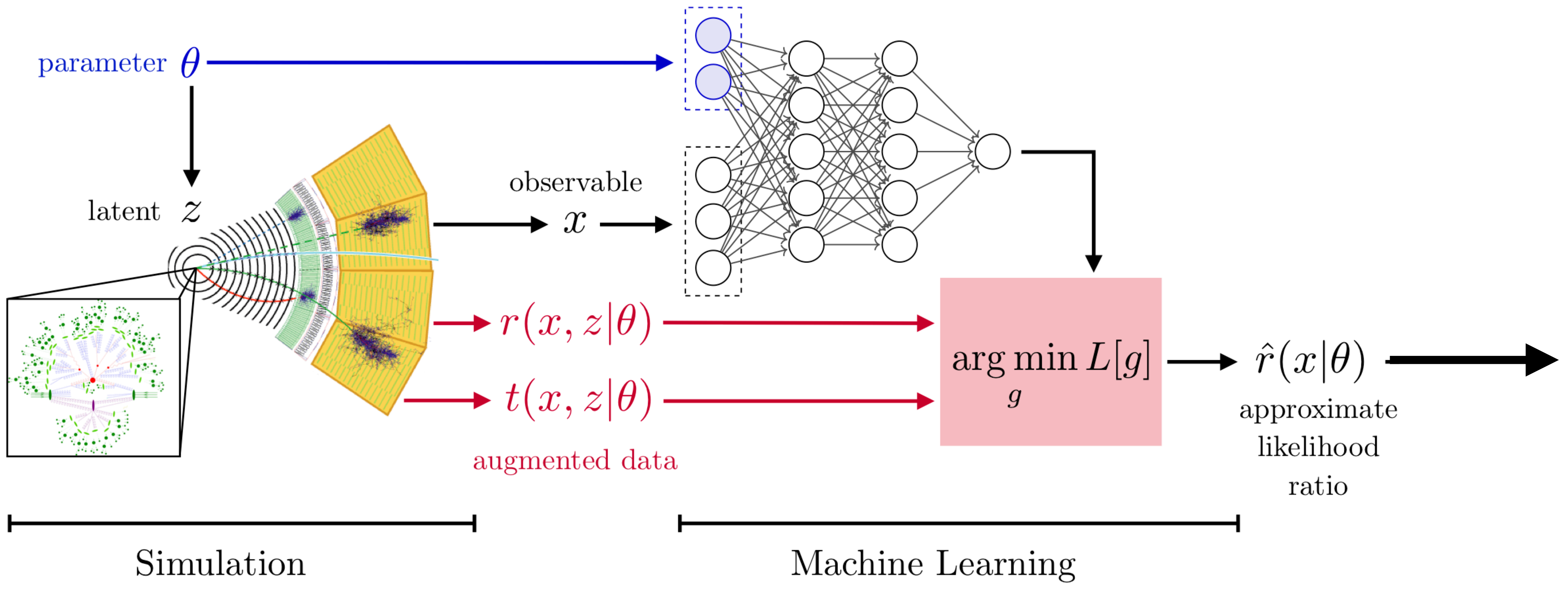
- Removes the need for hand-engineered summary statistics that sacrifice power

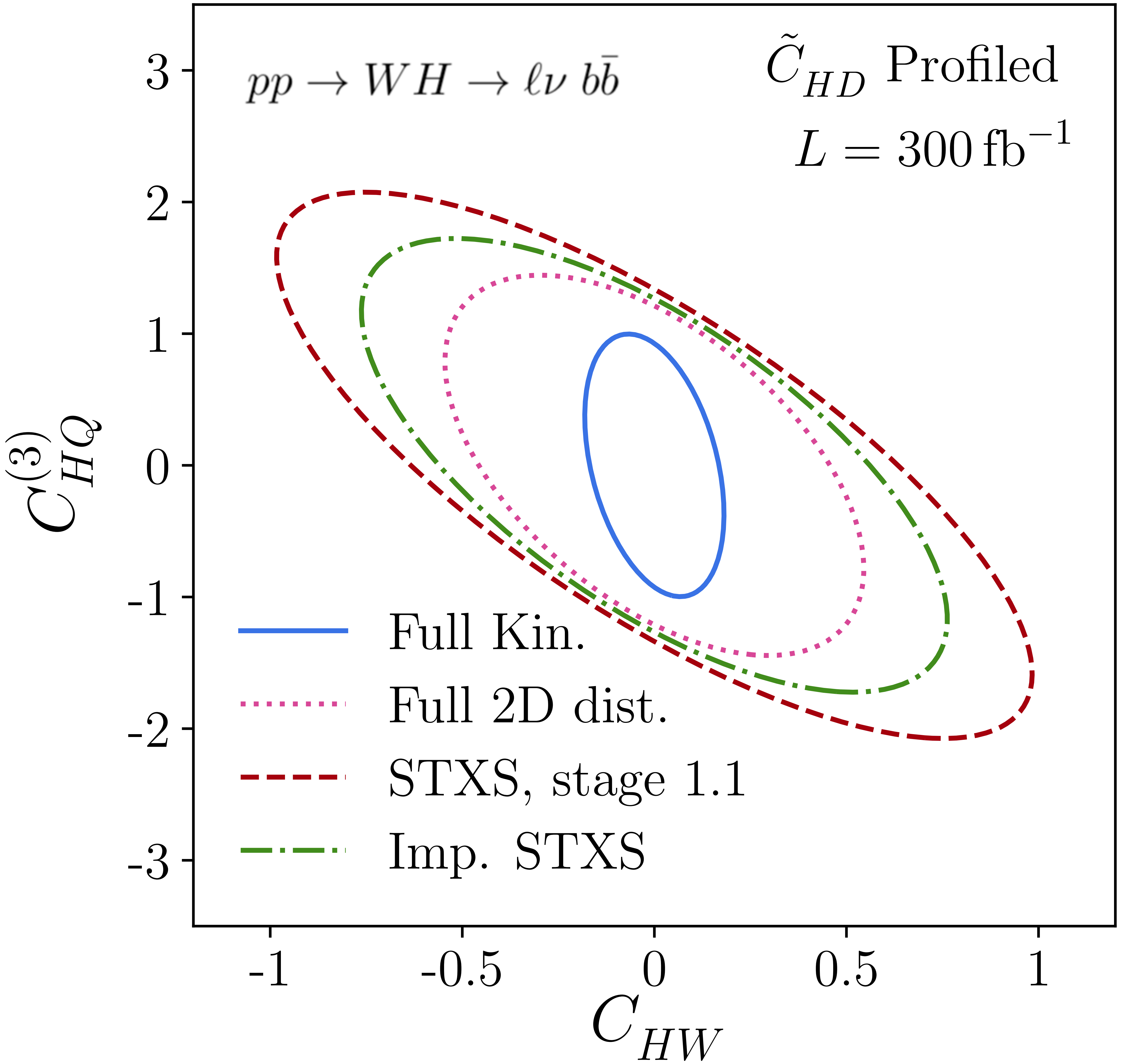


Impact on Studies of The Higgs Boson

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

Equivalent increasing data collected by LHC by several factors





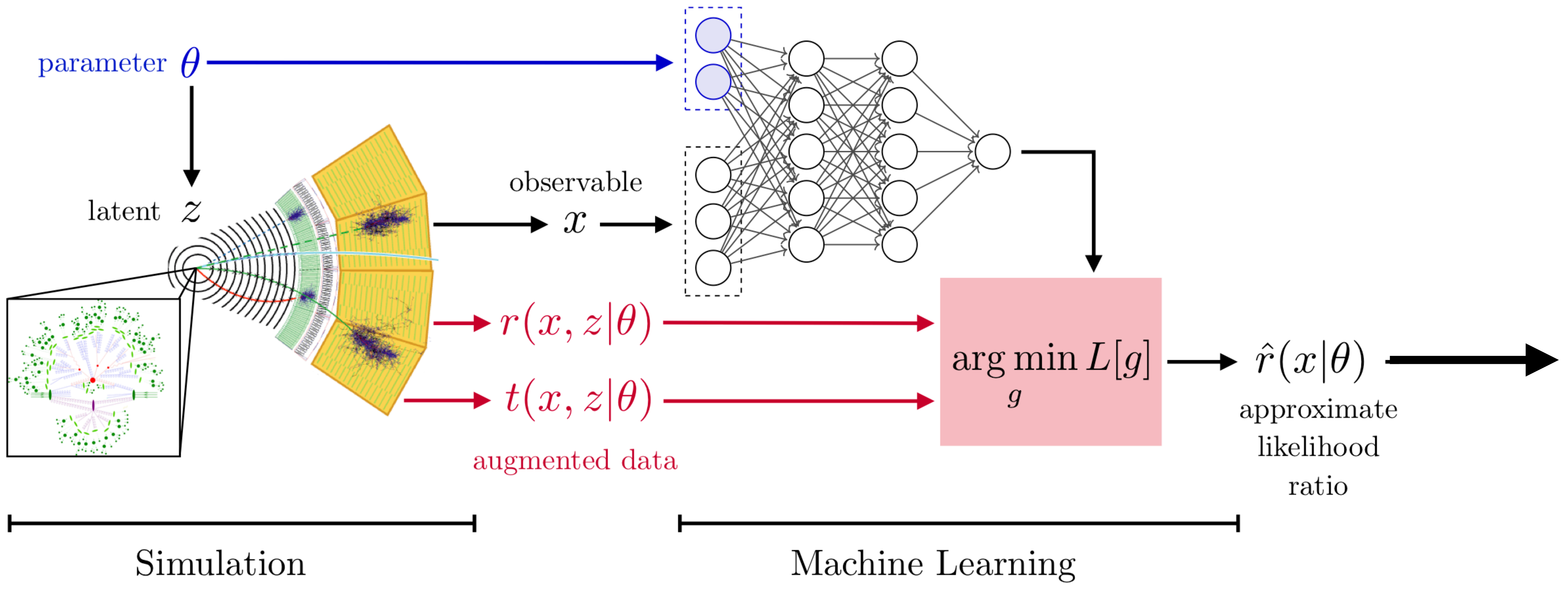
[J. Brehmer, S. Dawson, S. Homiller, F. Kling, T. Plehn 1908.06980]

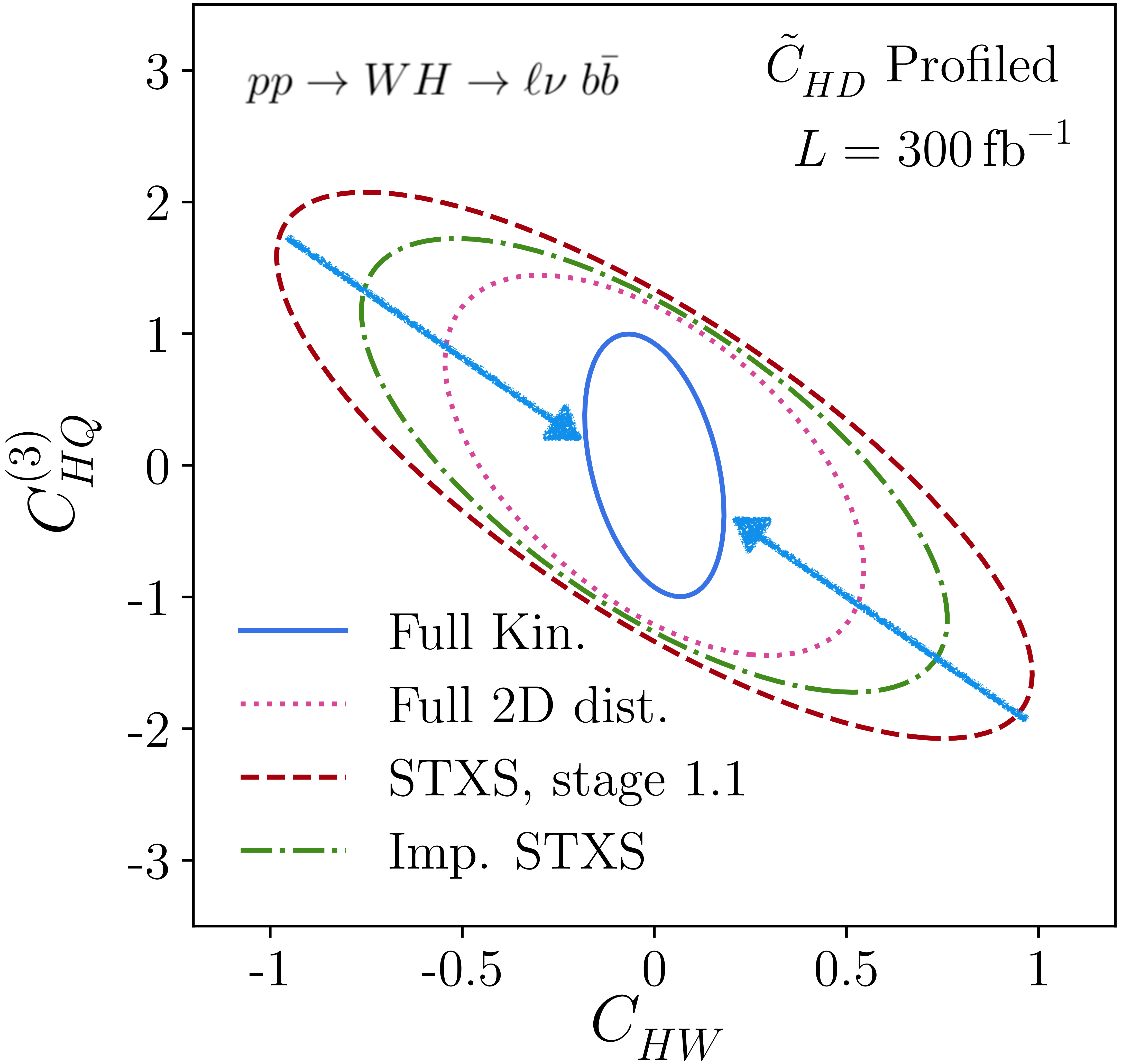
[J. Brehmer, F. Kling, I. Espejo, K. Cranmer 1907.10621]

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

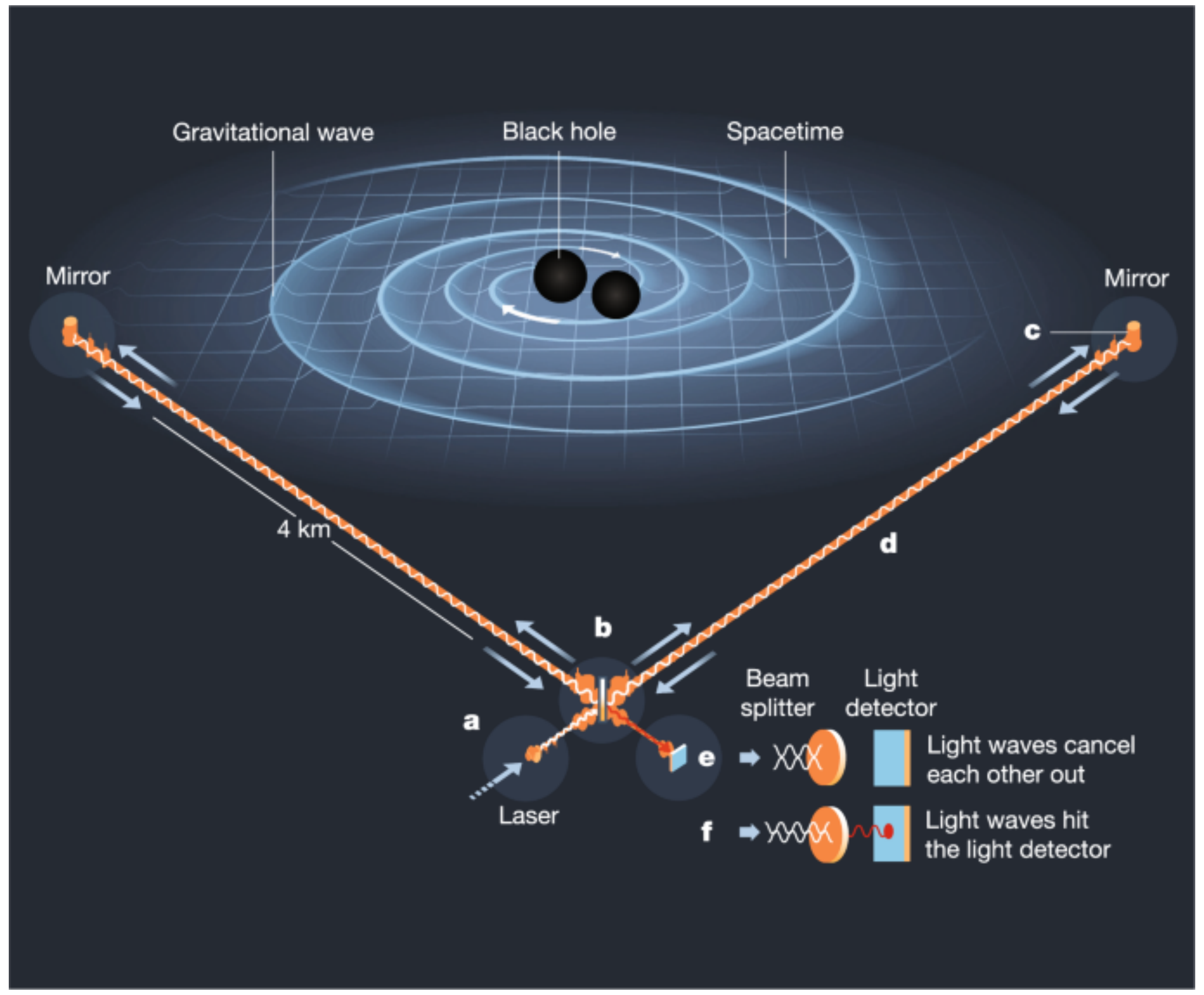


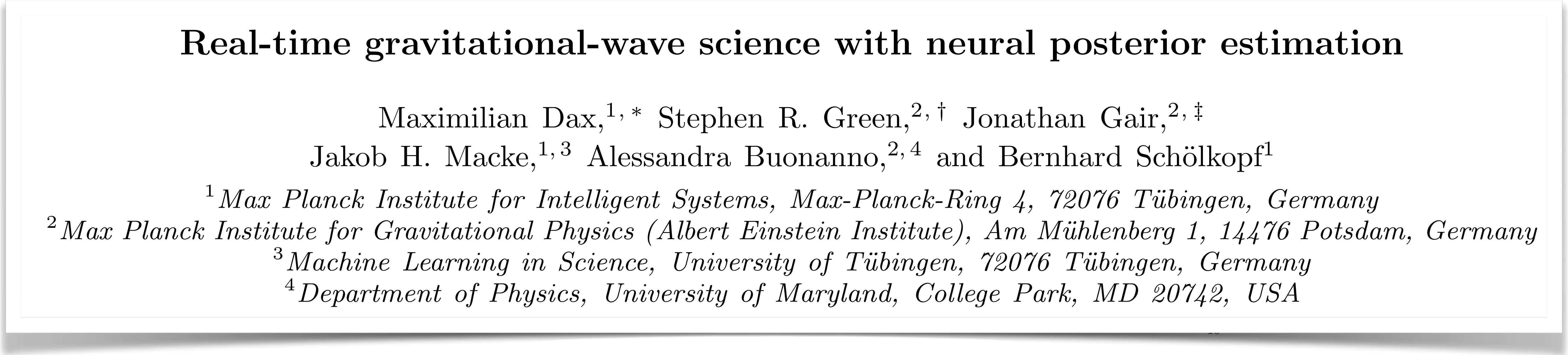


[J. Brehmer, S. Dawson, S. Homiller, F. Kling, T. Plehn 1908.06980]

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

Gravitational Wave Astronomy

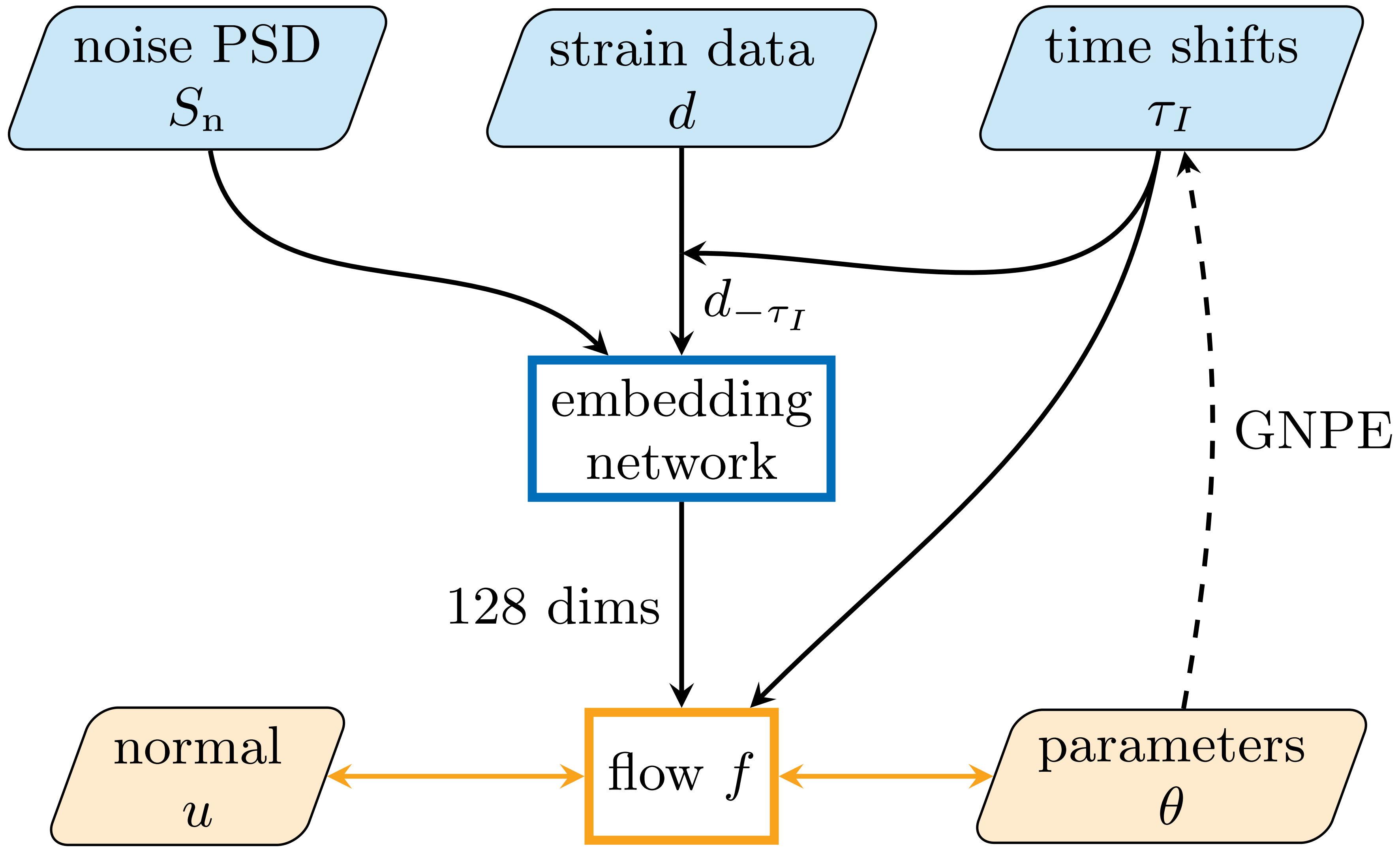




Real-time gravitational-wave science with neural posterior estimation

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*



noise PSD S_n

strain data d

time shifts τ_I

$d_{-\tau_I}$

embedding network

GNPE

128 dims

normal u

flow f

parameters θ

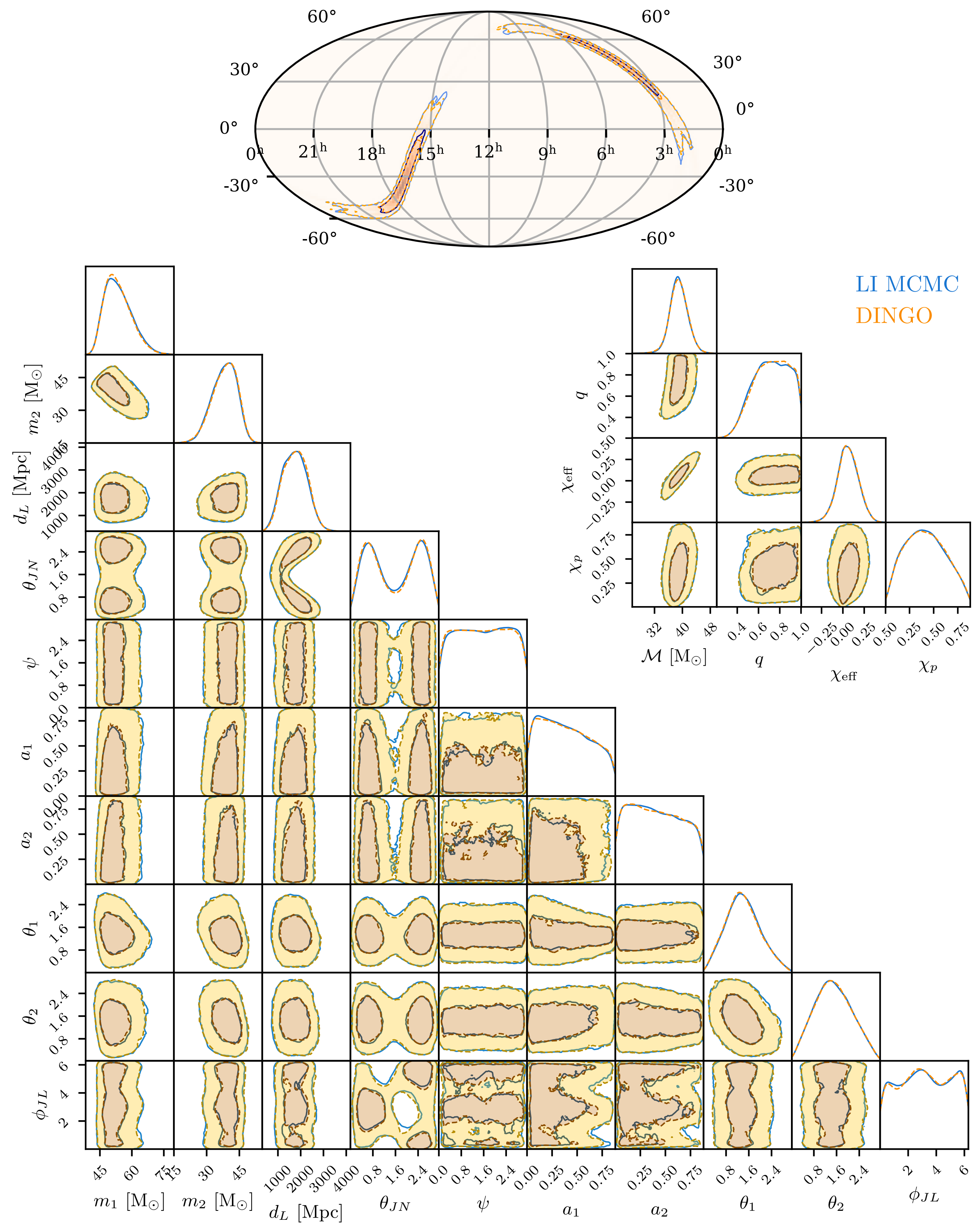
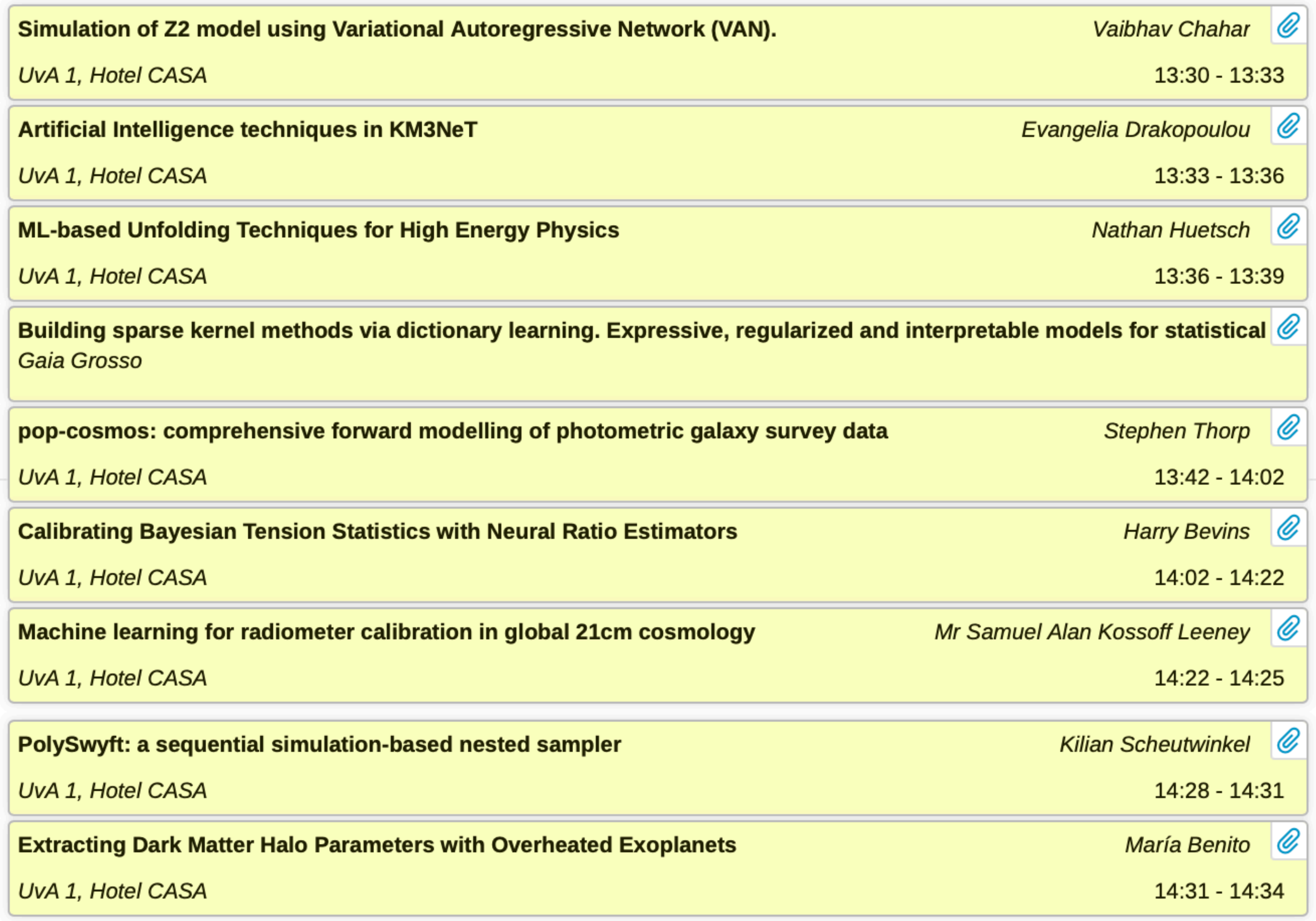


Figure 15. GW170823.

SBI at EuCAIFCon 2024

Simulation-based Inference was well represented at EuCAIFCon!



Simulation of Z2 model using Variational Autoregressive Network (VAN). *Vaibhav Chahar*
UvA 1, Hotel CASA 13:30 - 13:33

Artificial Intelligence techniques in KM3NeT *Evangelia Drakopoulou*
UvA 1, Hotel CASA 13:33 - 13:36

ML-based Unfolding Techniques for High Energy Physics *Nathan Huetsch*
UvA 1, Hotel CASA 13:36 - 13:39

Building sparse kernel methods via dictionary learning. Expressive, regularized and interpretable models for statistical
Gaia Grosso

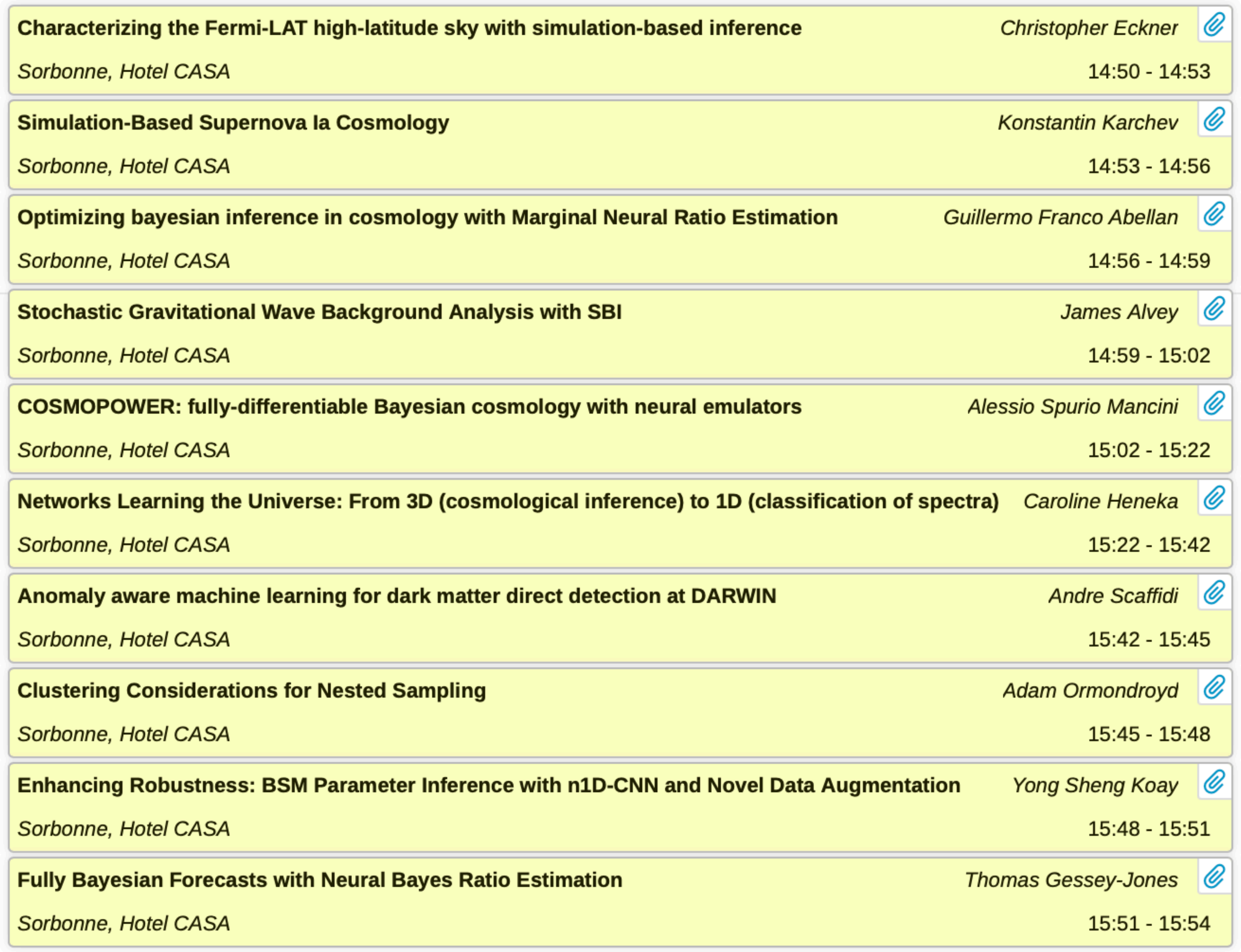
pop-cosmos: comprehensive forward modelling of photometric galaxy survey data *Stephen Thorp*
UvA 1, Hotel CASA 13:42 - 14:02

Calibrating Bayesian Tension Statistics with Neural Ratio Estimators *Harry Bevins*
UvA 1, Hotel CASA 14:02 - 14:22

Machine learning for radiometer calibration in global 21cm cosmology *Mr Samuel Alan Kossoff Leeney*
UvA 1, Hotel CASA 14:22 - 14:25

PolySwyft: a sequential simulation-based nested sampler *Kilian Scheutwinkel*
UvA 1, Hotel CASA 14:28 - 14:31

Extracting Dark Matter Halo Parameters with Overheated Exoplanets *María Benito*
UvA 1, Hotel CASA 14:31 - 14:34



Characterizing the Fermi-LAT high-latitude sky with simulation-based inference *Christopher Eckner*
Sorbonne, Hotel CASA 14:50 - 14:53

Simulation-Based Supernova Ia Cosmology *Konstantin Karchev*
Sorbonne, Hotel CASA 14:53 - 14:56

Optimizing bayesian inference in cosmology with Marginal Neural Ratio Estimation *Guillermo Franco Abellan*
Sorbonne, Hotel CASA 14:56 - 14:59

Stochastic Gravitational Wave Background Analysis with SBI *James Alvey*
Sorbonne, Hotel CASA 14:59 - 15:02

COSMOPOWER: fully-differentiable Bayesian cosmology with neural emulators *Alessio Spurio Mancini*
Sorbonne, Hotel CASA 15:02 - 15:22

Networks Learning the Universe: From 3D (cosmological inference) to 1D (classification of spectra) *Caroline Heneka*
Sorbonne, Hotel CASA 15:22 - 15:42

Anomaly aware machine learning for dark matter direct detection at DARWIN *Andre Scaffidi*
Sorbonne, Hotel CASA 15:42 - 15:45

Clustering Considerations for Nested Sampling *Adam Ormondroyd*
Sorbonne, Hotel CASA 15:45 - 15:48

Enhancing Robustness: BSM Parameter Inference with n1D-CNN and Novel Data Augmentation *Yong Sheng Koay*
Sorbonne, Hotel CASA 15:48 - 15:51

Fully Bayesian Forecasts with Neural Bayes Ratio Estimation *Thomas Gessey-Jones*
Sorbonne, Hotel CASA 15:51 - 15:54

Analyzing ML-enabled Full Population Model for Galaxy SEDs with Unsupervised Learning and Mutual Information
Dr Sinan Deger

Convolutional neural network search for long-duration transient gravitational waves from glitching pulsars
Rodrigo Tenorio

Tuning neural posterior estimation for gravitational wave inference *Alex Kolmus*
Oxford, Hotel CASA 16:06 - 16:09

Normalising flows for dense matter equation of state inference from gravitational wave observations of neutron star me
Jessica Irwin

A Strong Gravitational Lens Is Worth a Thousand Dark Matter Halos: Inference on Small-Scale Structure Using Sequent
Sebastian Wagner-Carena

Optimal, fast, and robust inference of reionization-era cosmology with the 21cmPIE-INN *Benedikt Schosser*
Oxford, Hotel CASA 16:49 - 16:52

Simulation Based Inference from the CD-EoR 21-cm signal *Anchal Saxena*
Oxford, Hotel CASA 16:52 - 16:55

Flexible conditional normalizing flow distributions over manifolds: the jammy-flows toolkit *Dr Thorsten Glüsenkamp*
Oxford, Hotel CASA 16:55 - 16:58

A deep learning method for the trajectory reconstruction of gamma rays with the DAMPE space mission
Parzival Nussbaum

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

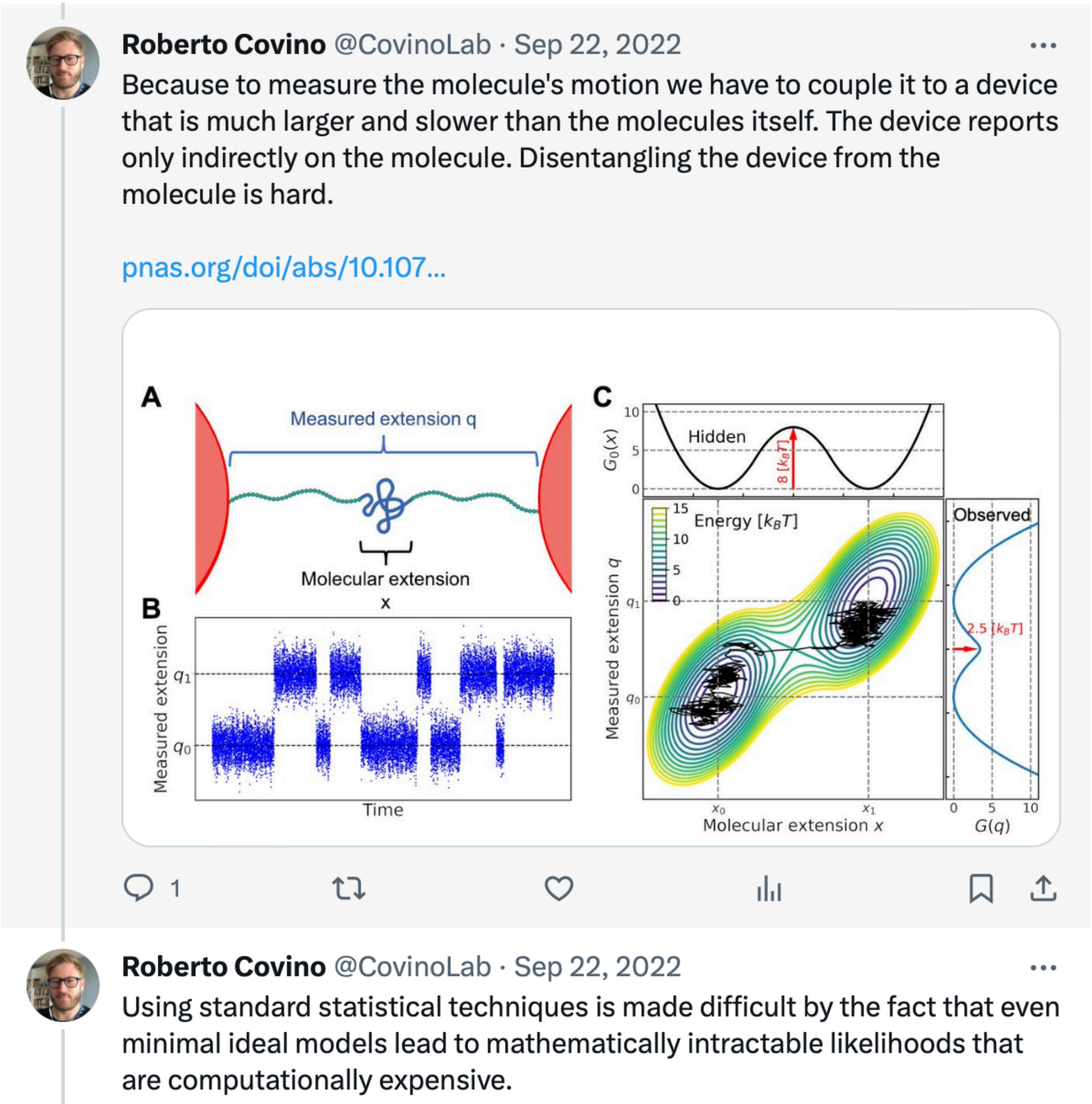
Summary talks: Cosmology and AI (Benjamin Wandelt) *David Rousseau*
UvA 2-3-4, Hotel CASA 15:00 - 15:40

Biophysical Journal Supports open access

ABSTRACT VOLUME 122, ISSUE 3, SUPPLEMENT 1, 140A, FEBRUARY 10, 2023

Simulation-based inference of single-molecule force spectroscopy

Roberto Covino • Pilar Cossio • Lars Dingeldein



Roberto Covino @CovinoLab · Sep 22, 2022

Because to measure the molecule's motion we have to couple it to a device that is much larger and slower than the molecules itself. The device reports only indirectly on the molecule. Disentangling the device from the molecule is hard.

pnas.org/doi/abs/10.107...

1

Roberto Covino @CovinoLab · Sep 22, 2022

Using standard statistical techniques is made difficult by the fact that even minimal ideal models lead to mathematically intractable likelihoods that are computationally expensive.

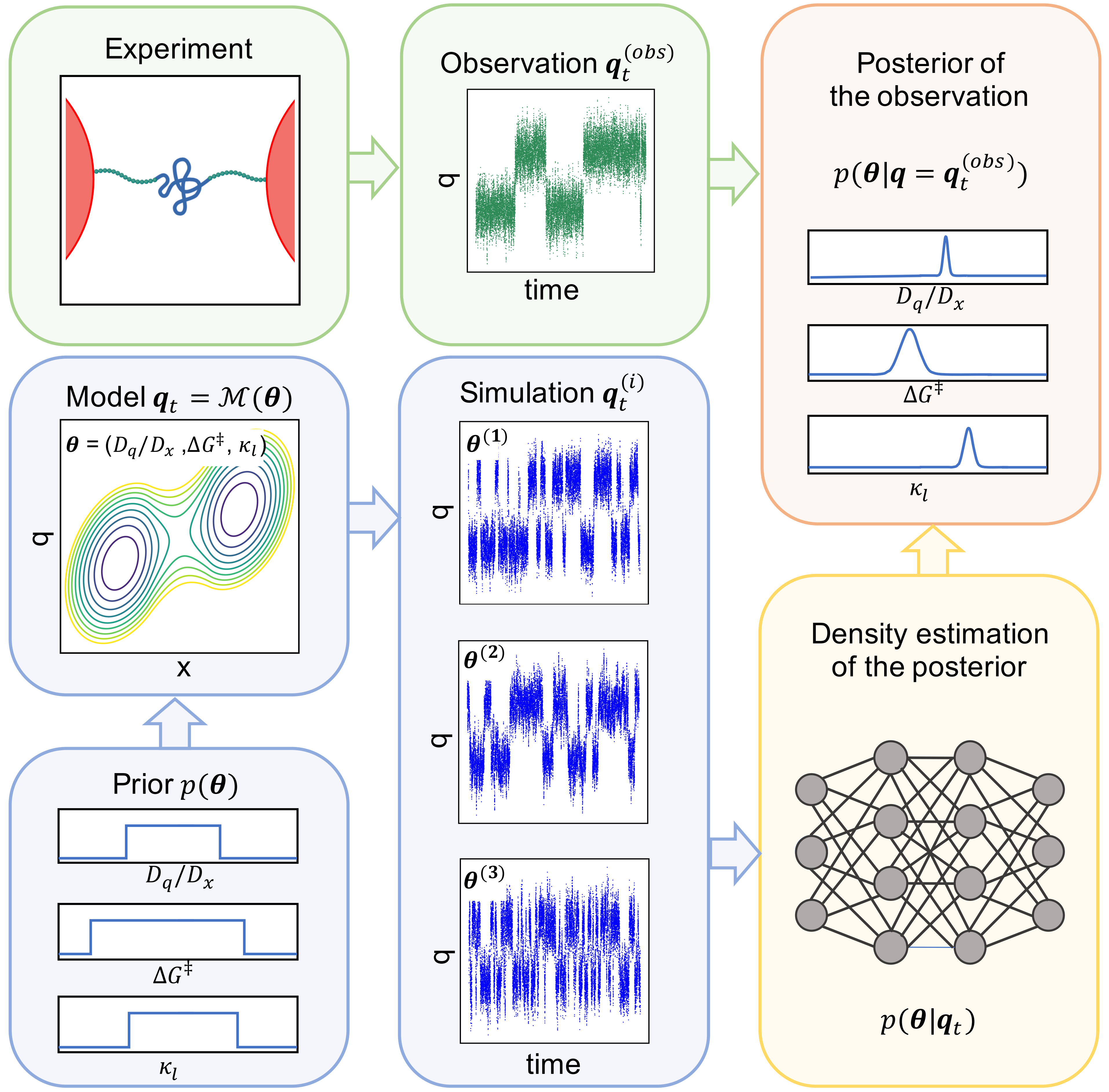
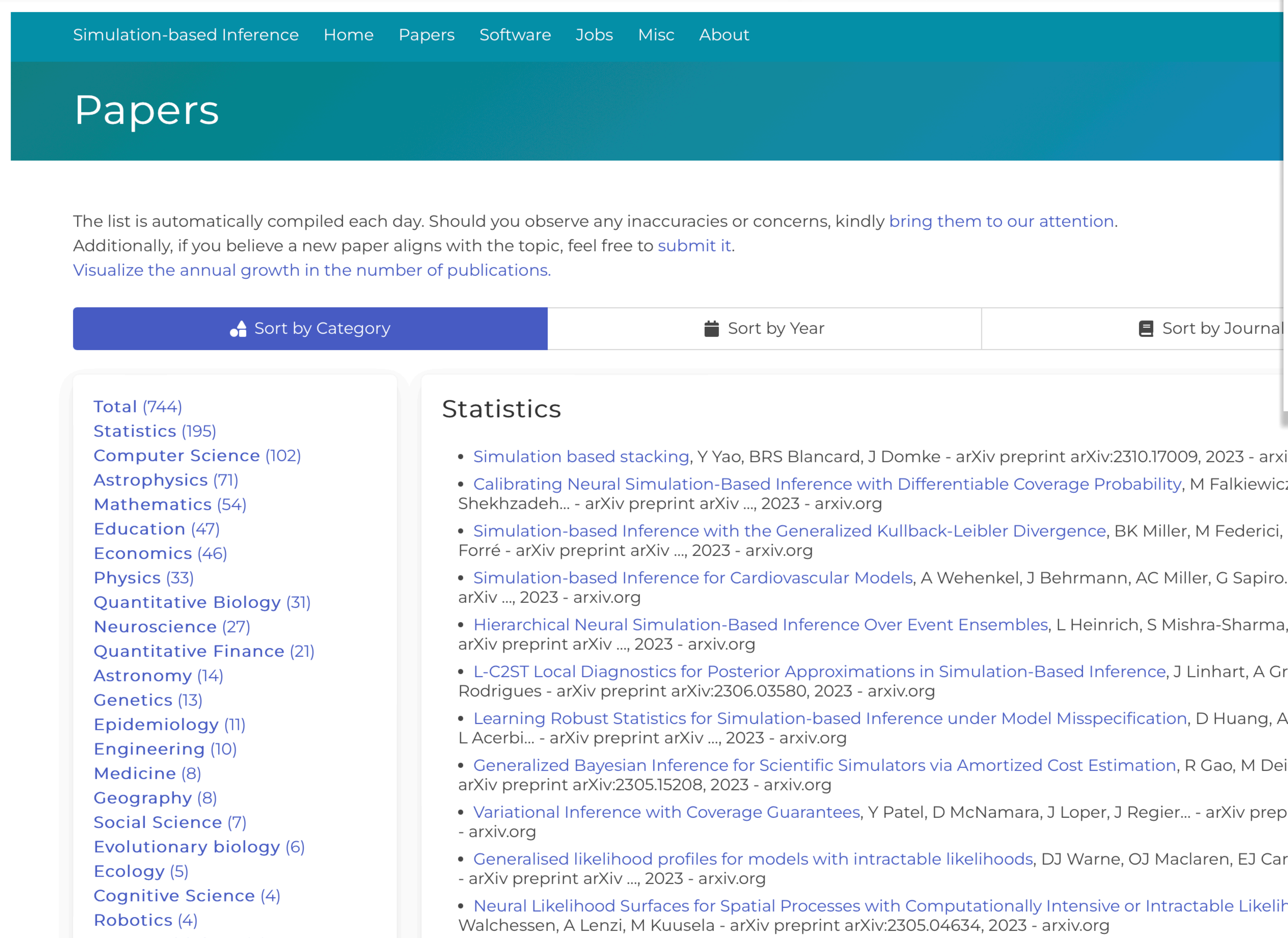


Fig. 2. **Workflow of Simulation-Based Inference of smFS experiments**. Given an experimental observation $q_t^{(obs)}$, we want to explain it with a parametric mechanistic model $\mathcal{M}(\theta)$ encoded in a simulator. The parameters θ are drawn from the prior distribution $p(\theta)$. The simulator takes parameters drawn from the prior and generates synthetic observations $q_t^{(i)}$. The synthetic observations and corresponding parameters are used to train a conditional density estimator to approximate the posterior $p(\theta|q_t)$. Evaluating the posterior, $p(\theta|q_t = q_t^{(obs)})$, we obtain the most plausible parameters explaining a given observation.

simulation-based-inference.org



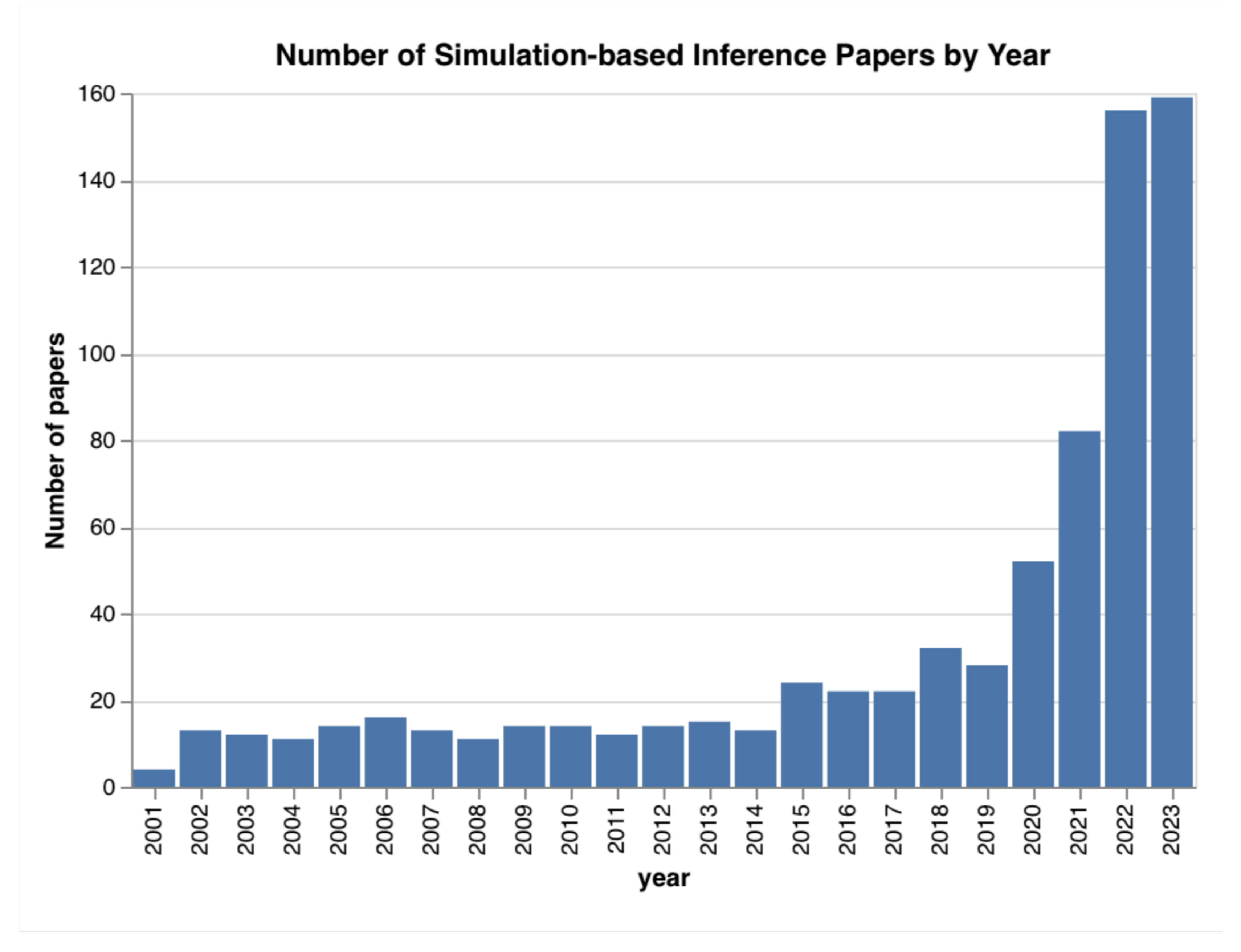
Papers

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.

- 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



Number of Simulation-based Inference Papers by Year

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

Computational Microscope

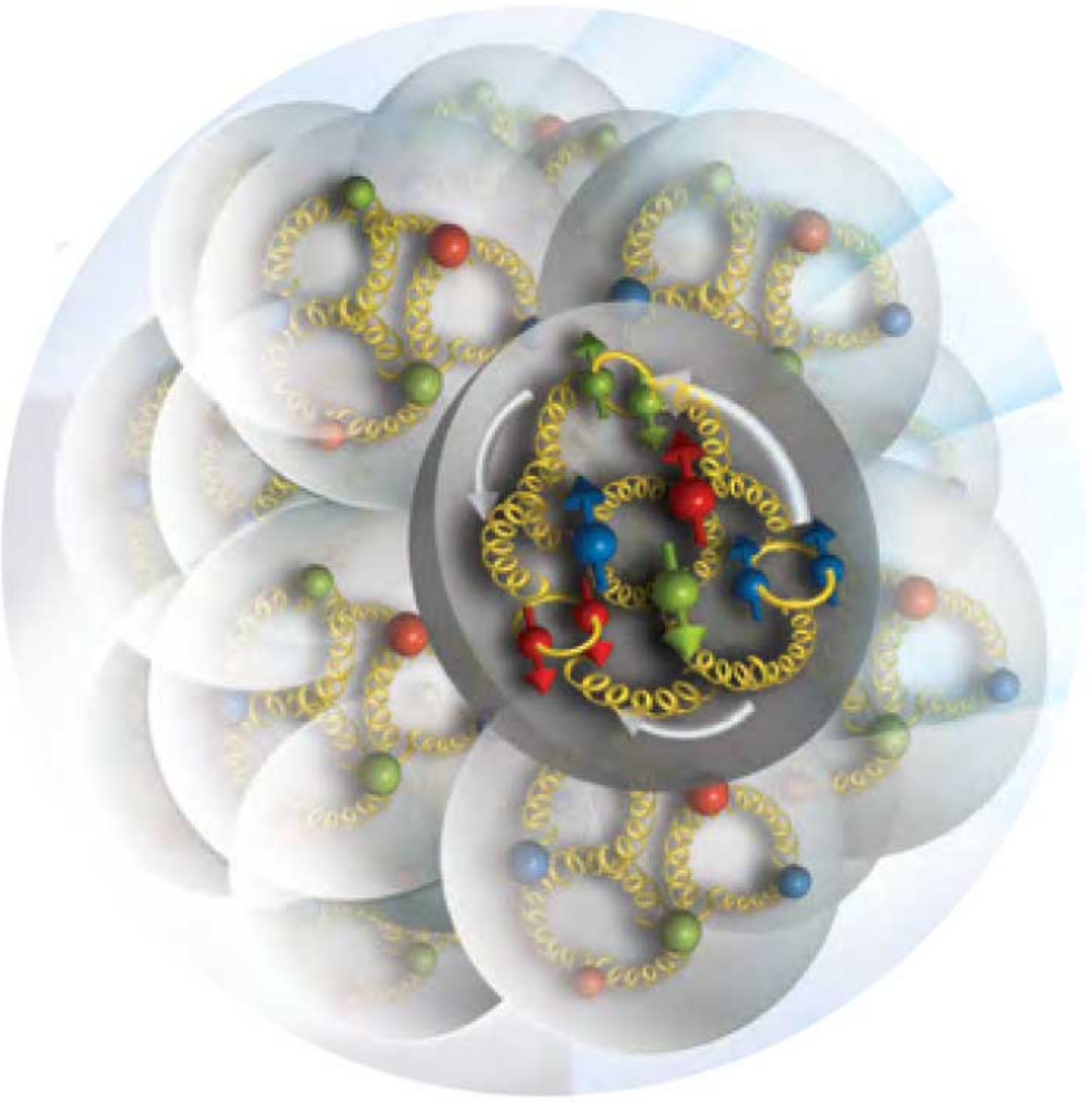
The strong force: Quantum Chromodynamics (QCD)

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

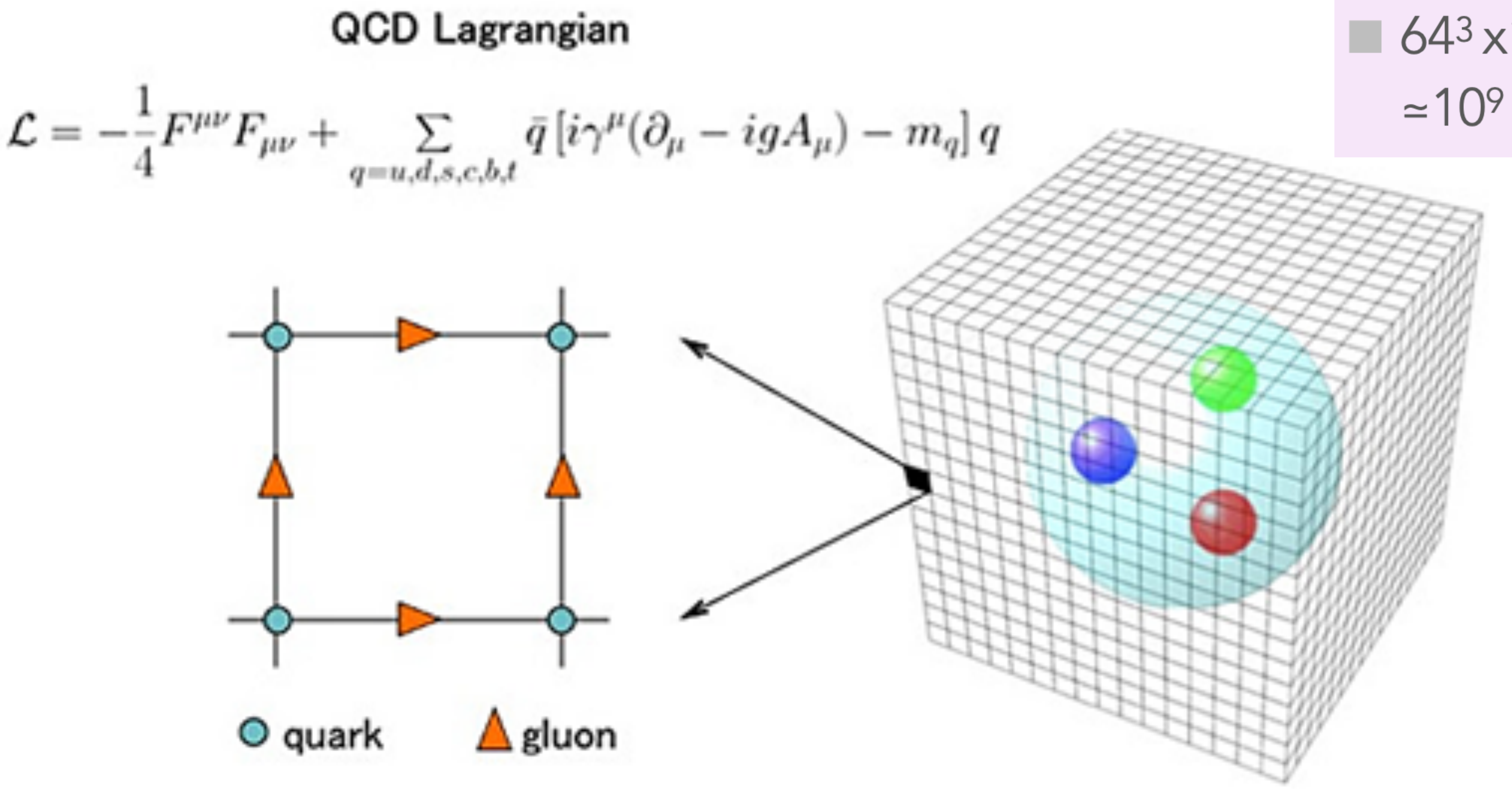
Emergent phenomena:
Quarks and gluons form protons, neutrons, etc.



Lattice Field Theory

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

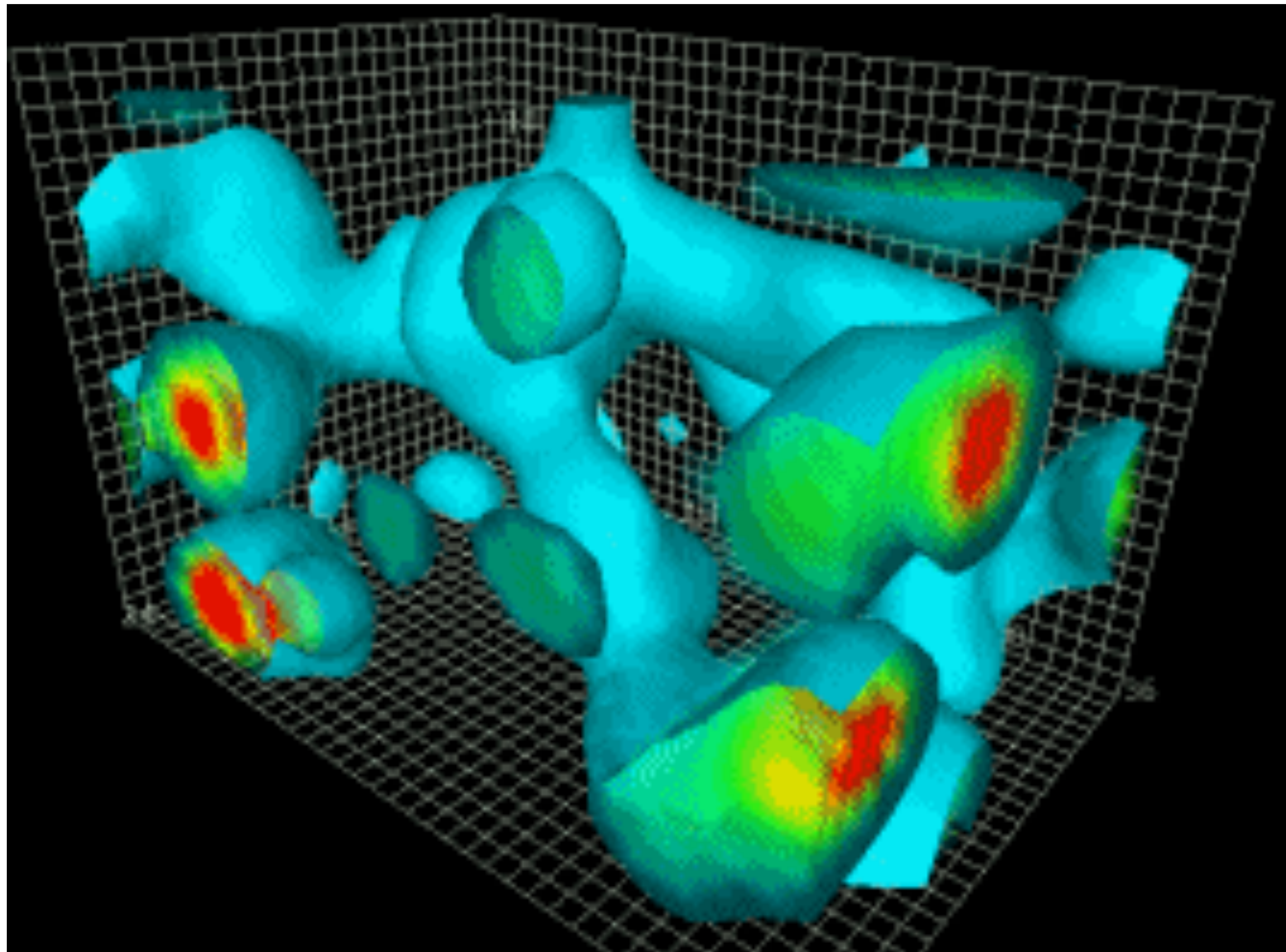
Each link on the lattice has data corresponding to the symmetry group of the theory. For the strong force (QCD) each link has a 3x3 unitary matrix.



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

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

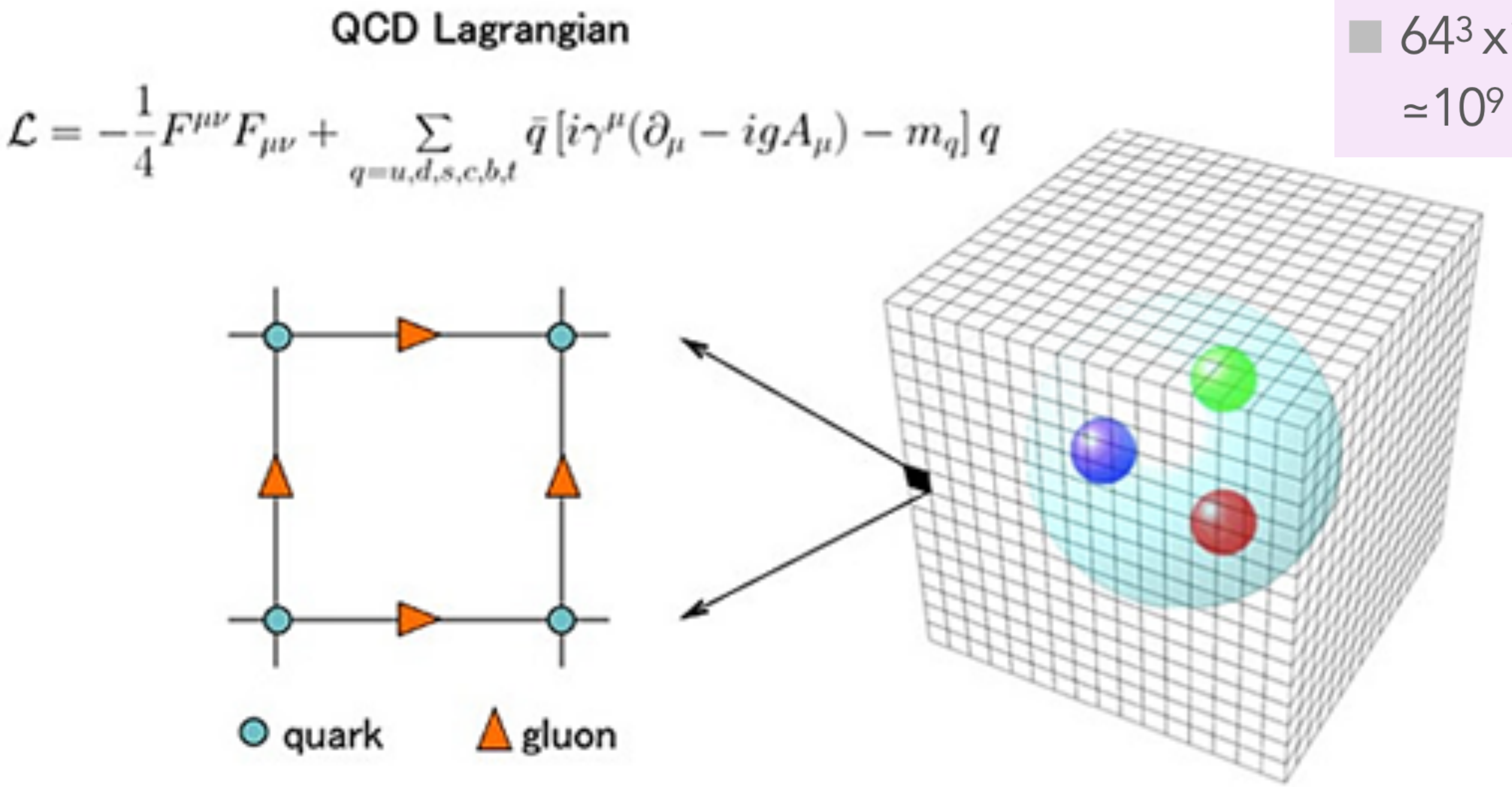
This animation is a single configuration of the lattice. Think of a 4-d image playing like a movie.



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.



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

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

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

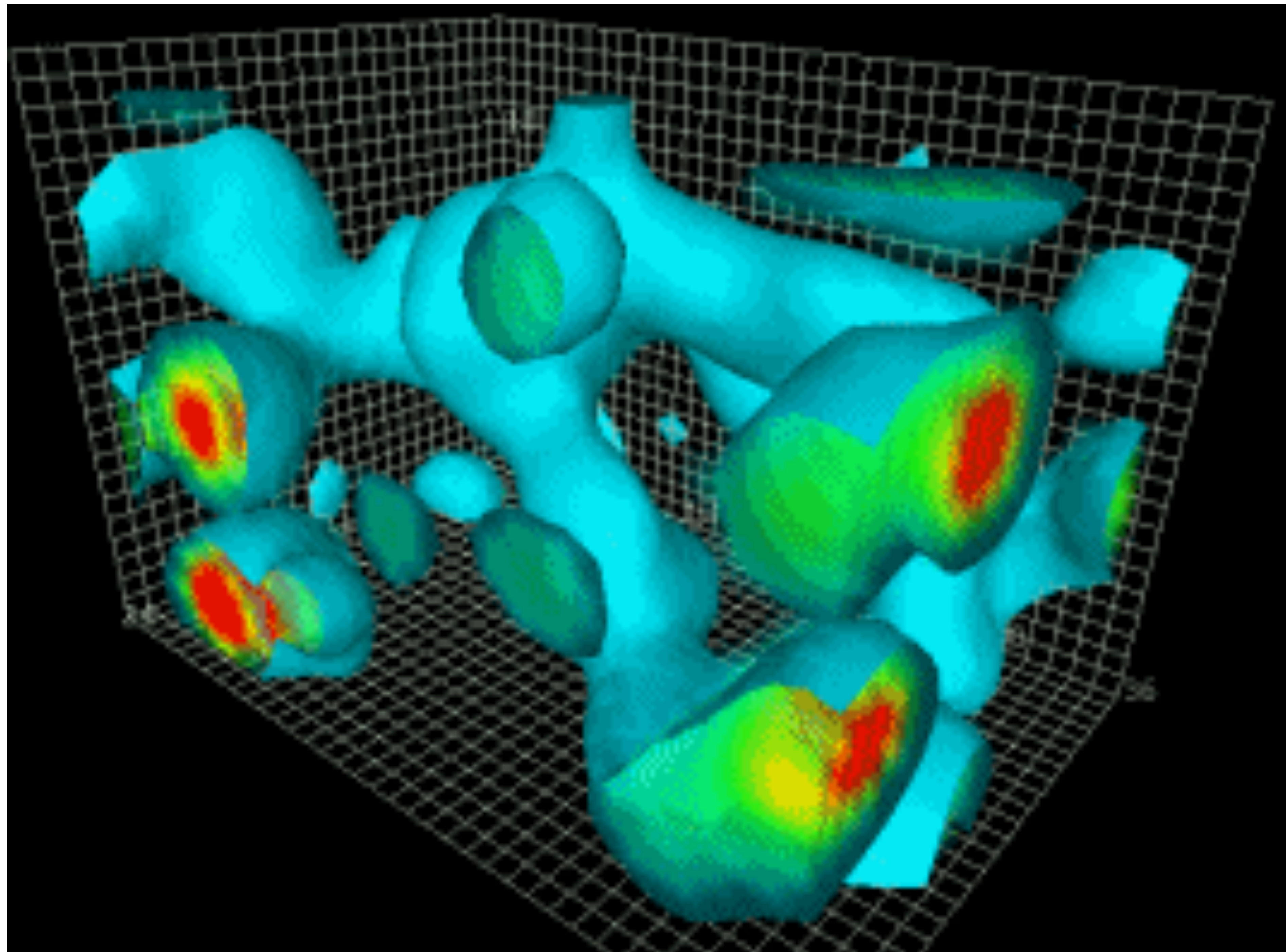
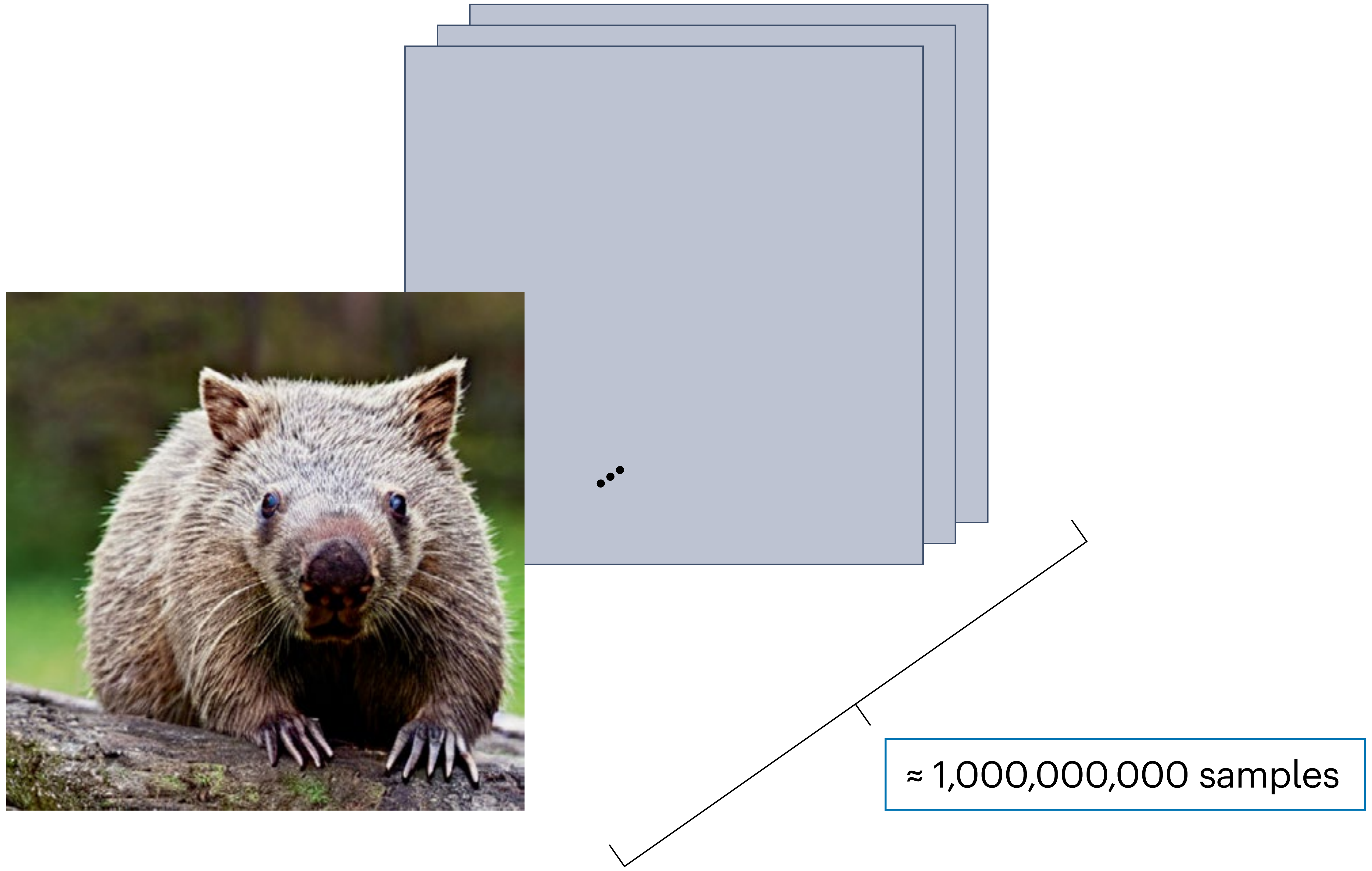


Image vs. Lattice Quantum Fields

Image generation



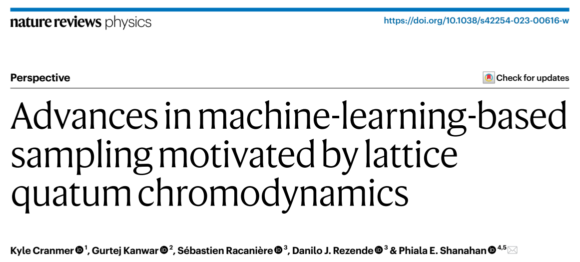
≈ 1,000,000,000 samples

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

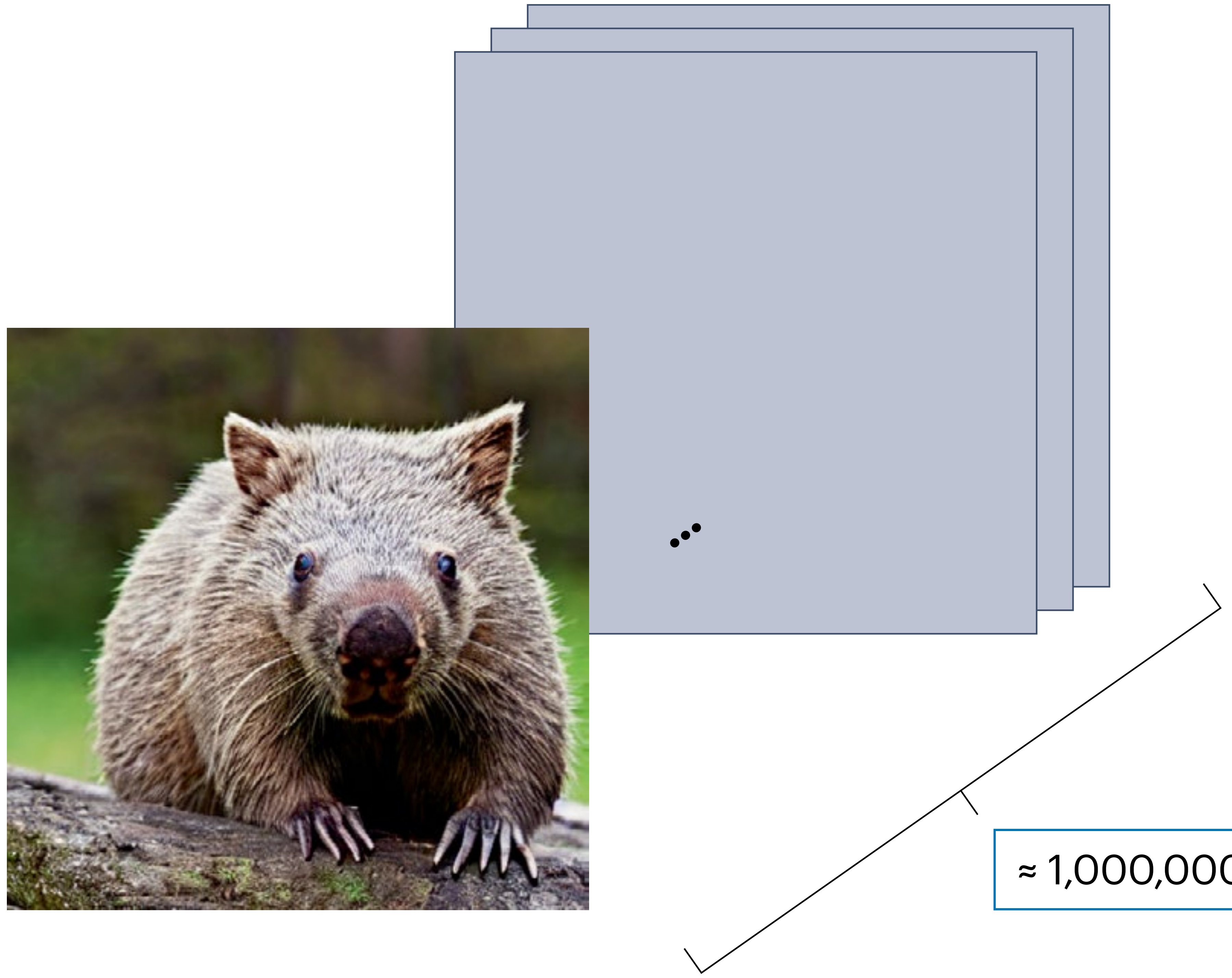
Perspective

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]

Image generation



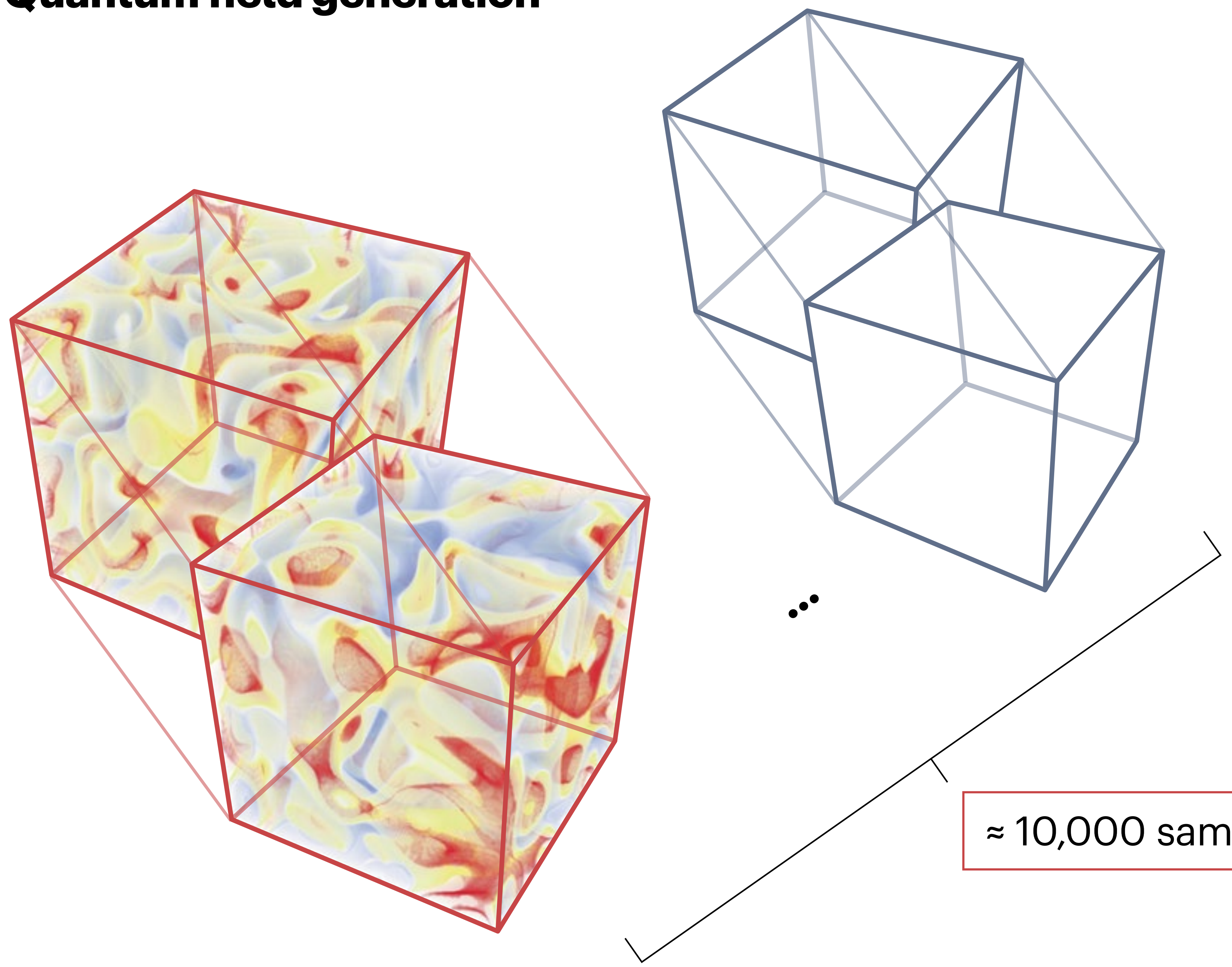
≈ 1,000,000,000 samples

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



≈ 10,000 samples

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

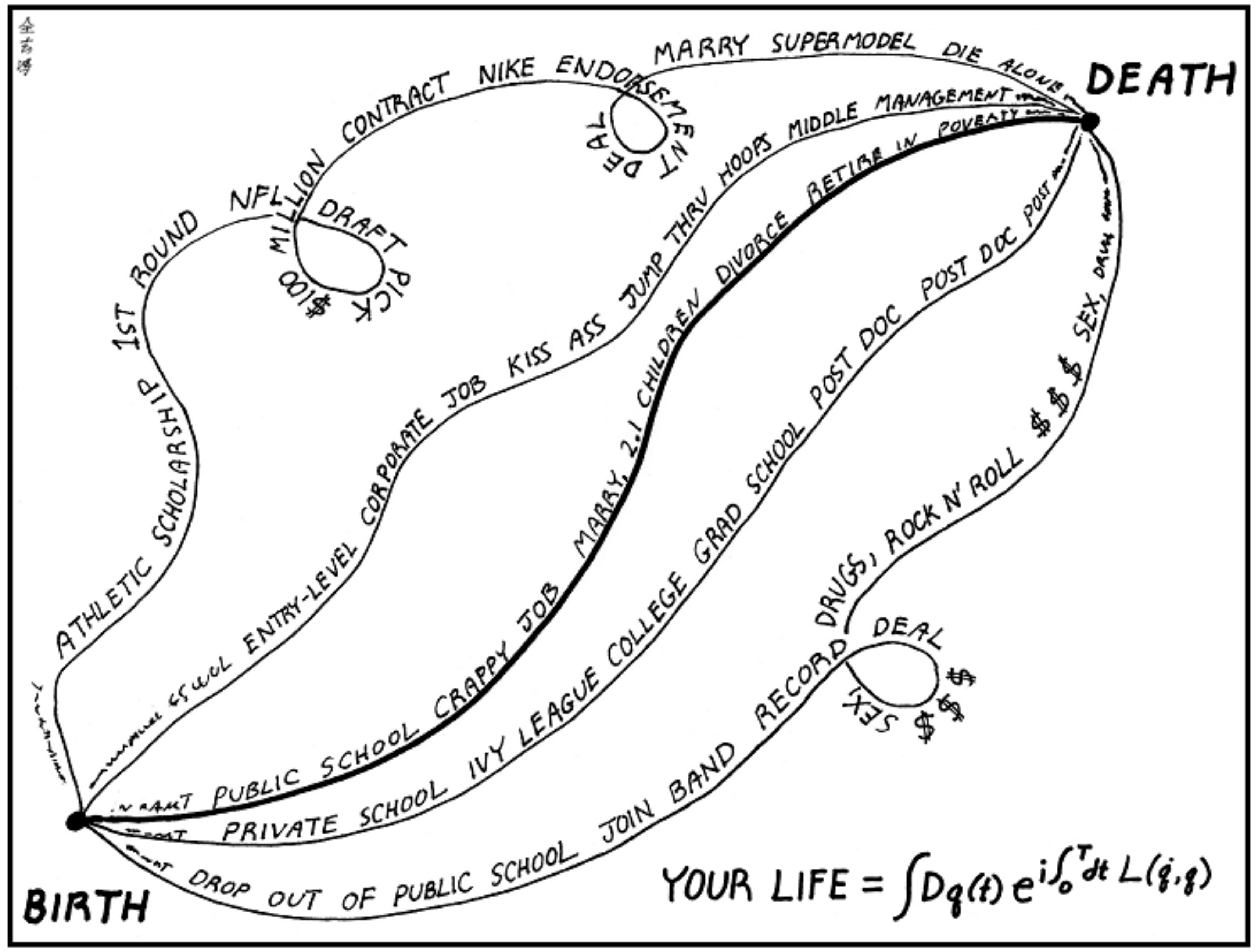
Target
Objective distribution $p(U) = e^{-S(U)}/Z$

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

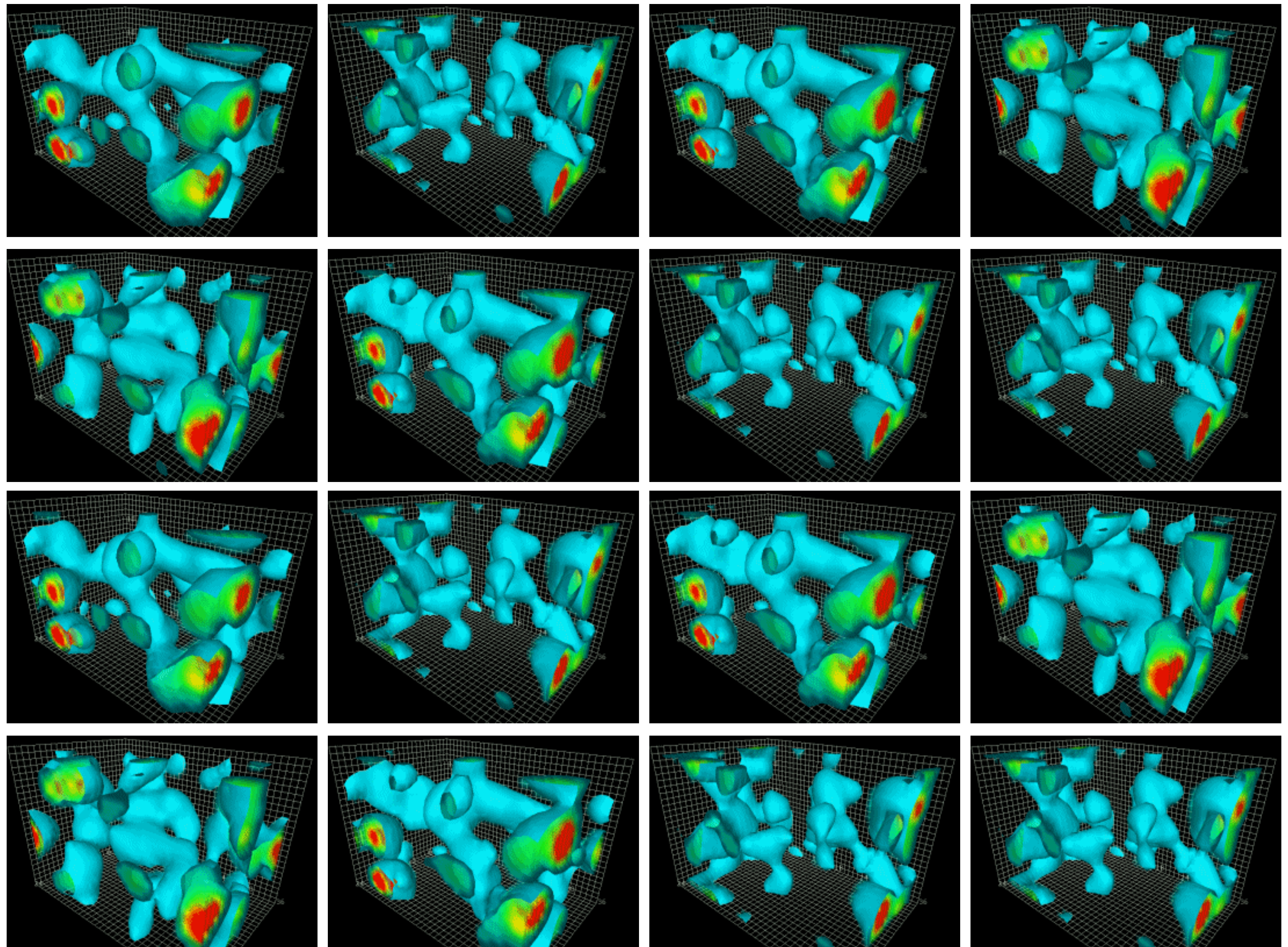
Distribution over configurations

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.



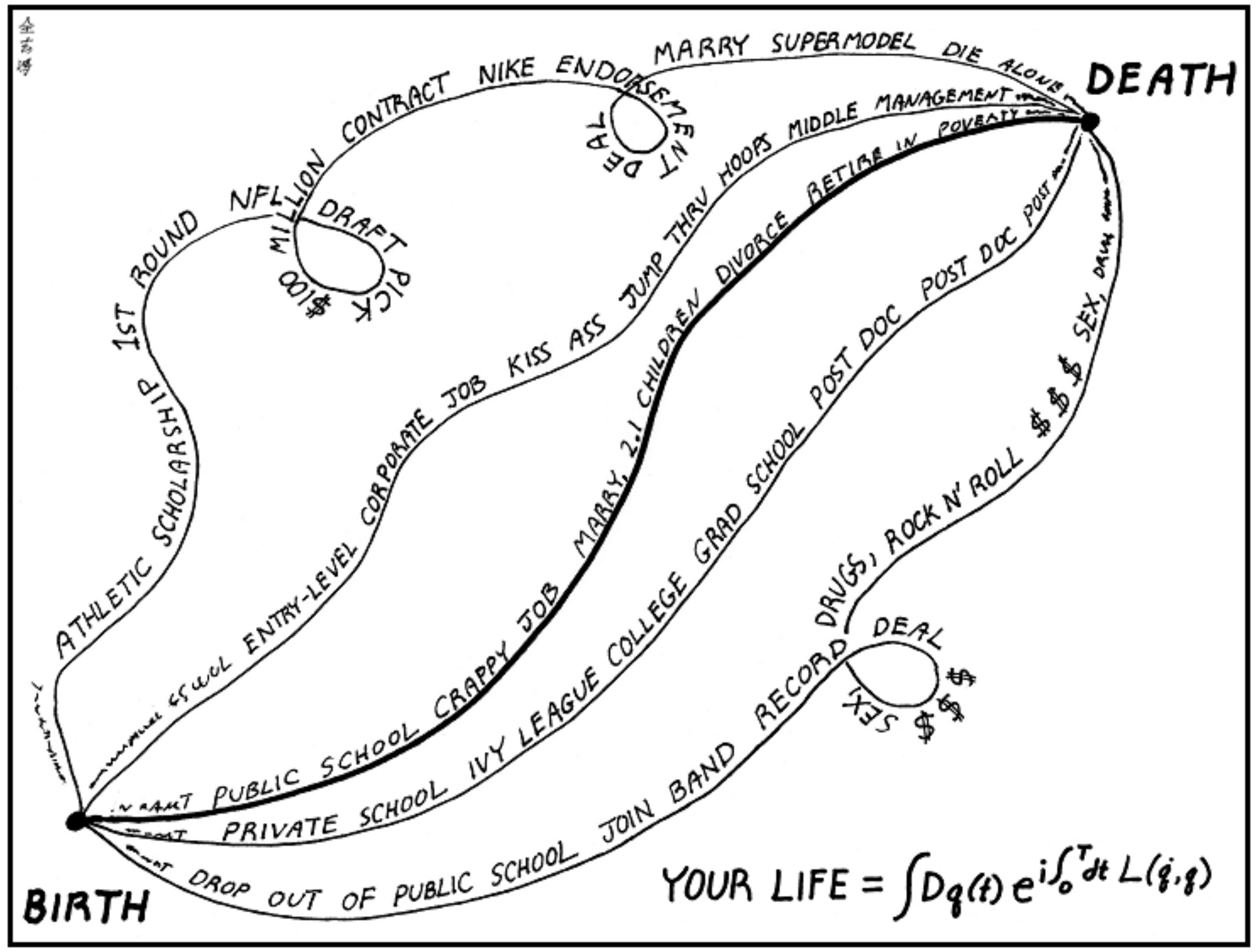
The Path Integral Formulation of Your Life



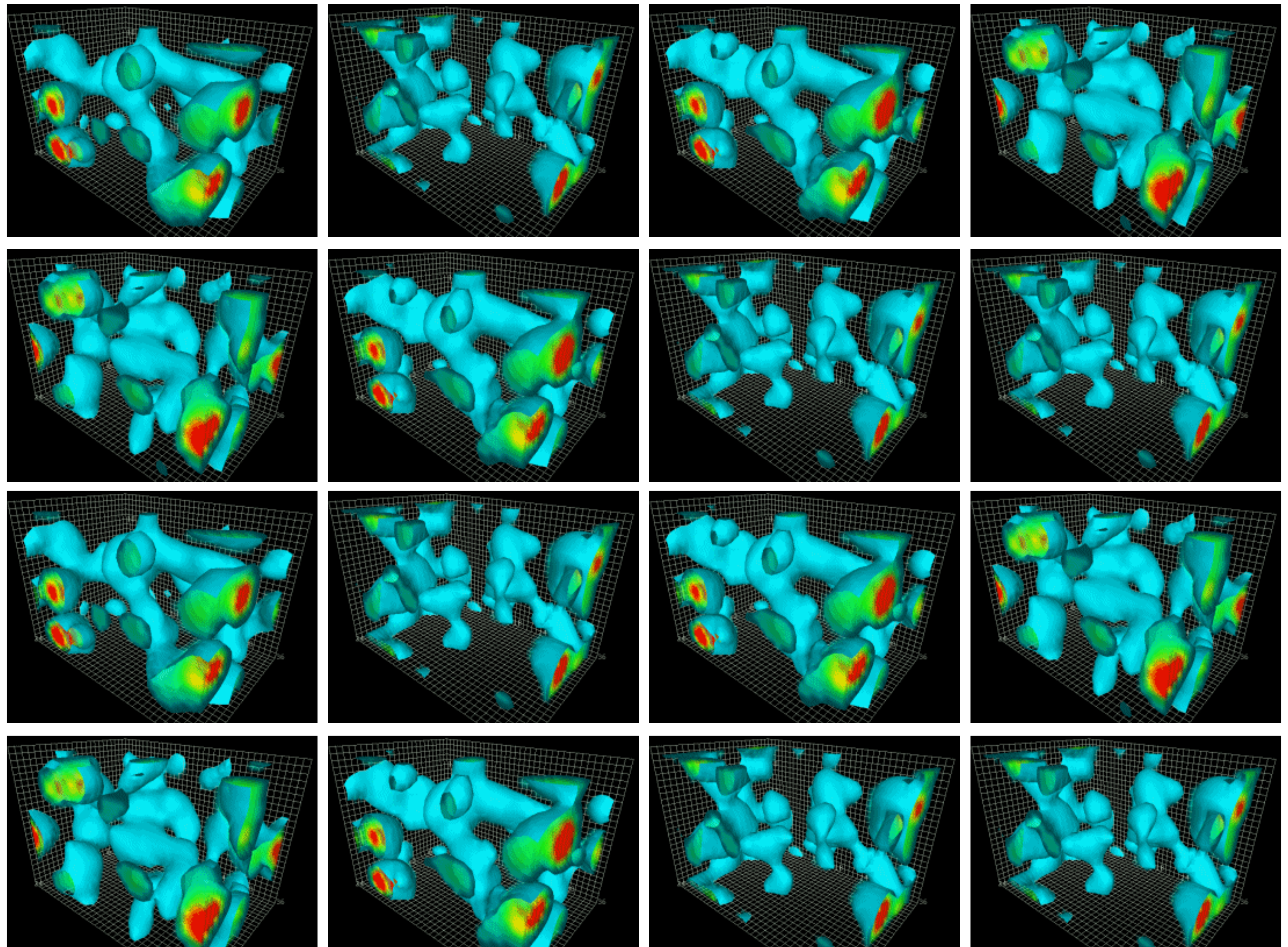
Distribution over configurations

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.

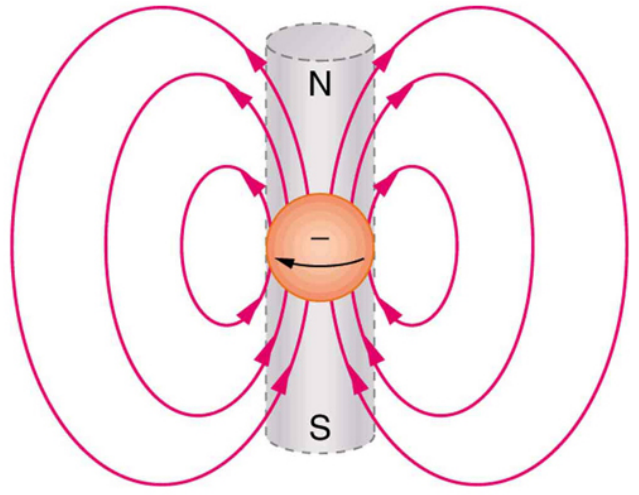


The Path Integral Formulation of Your Life



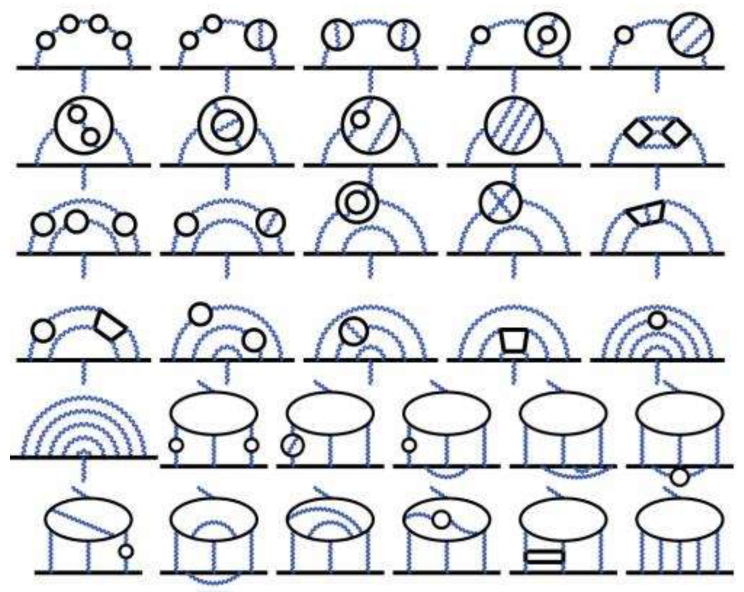
Predictions are taken seriously

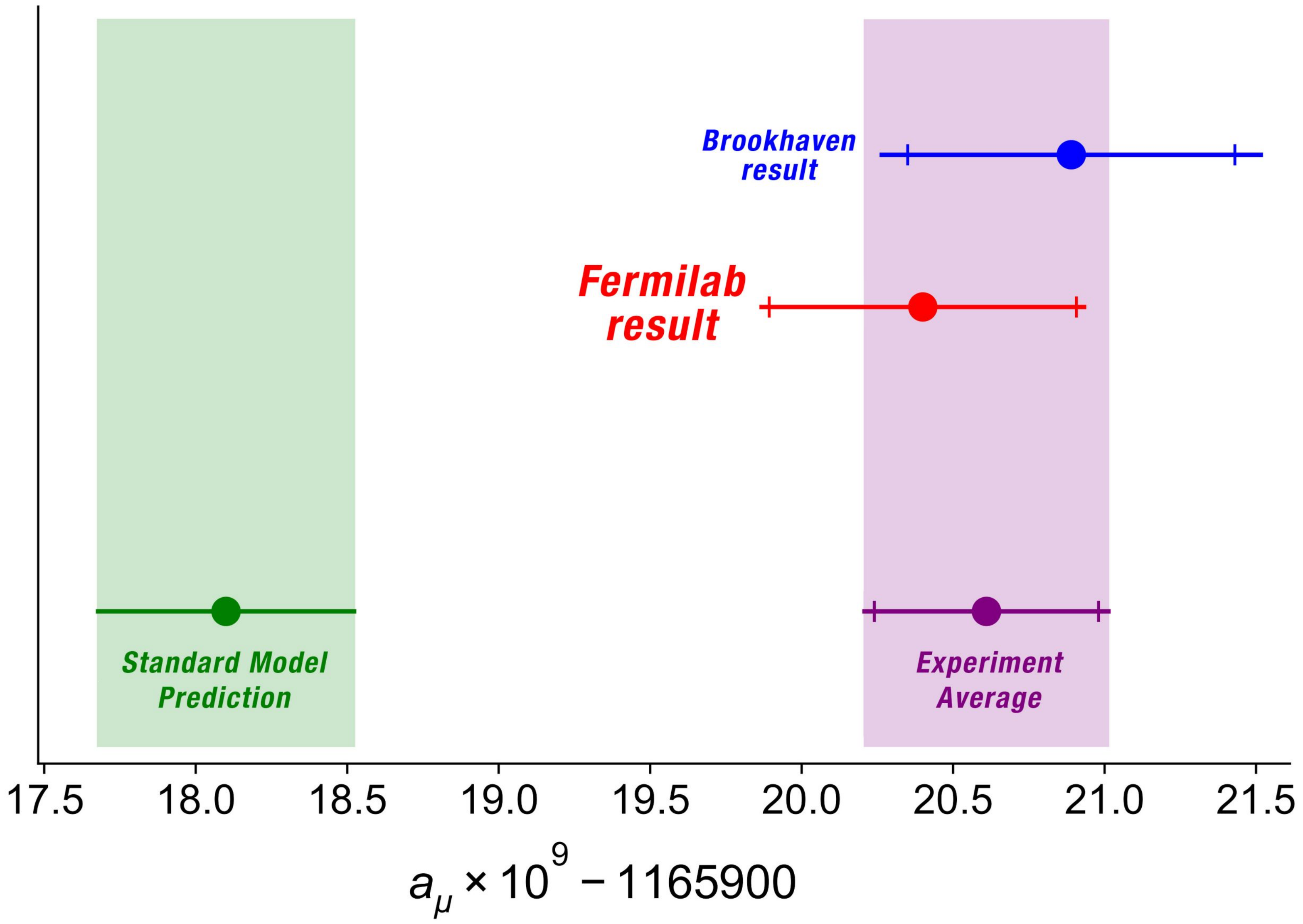
Magnetic moment of the electron:
(torque an electron feels in a magnetic field) $a_e = (g-2)/2$

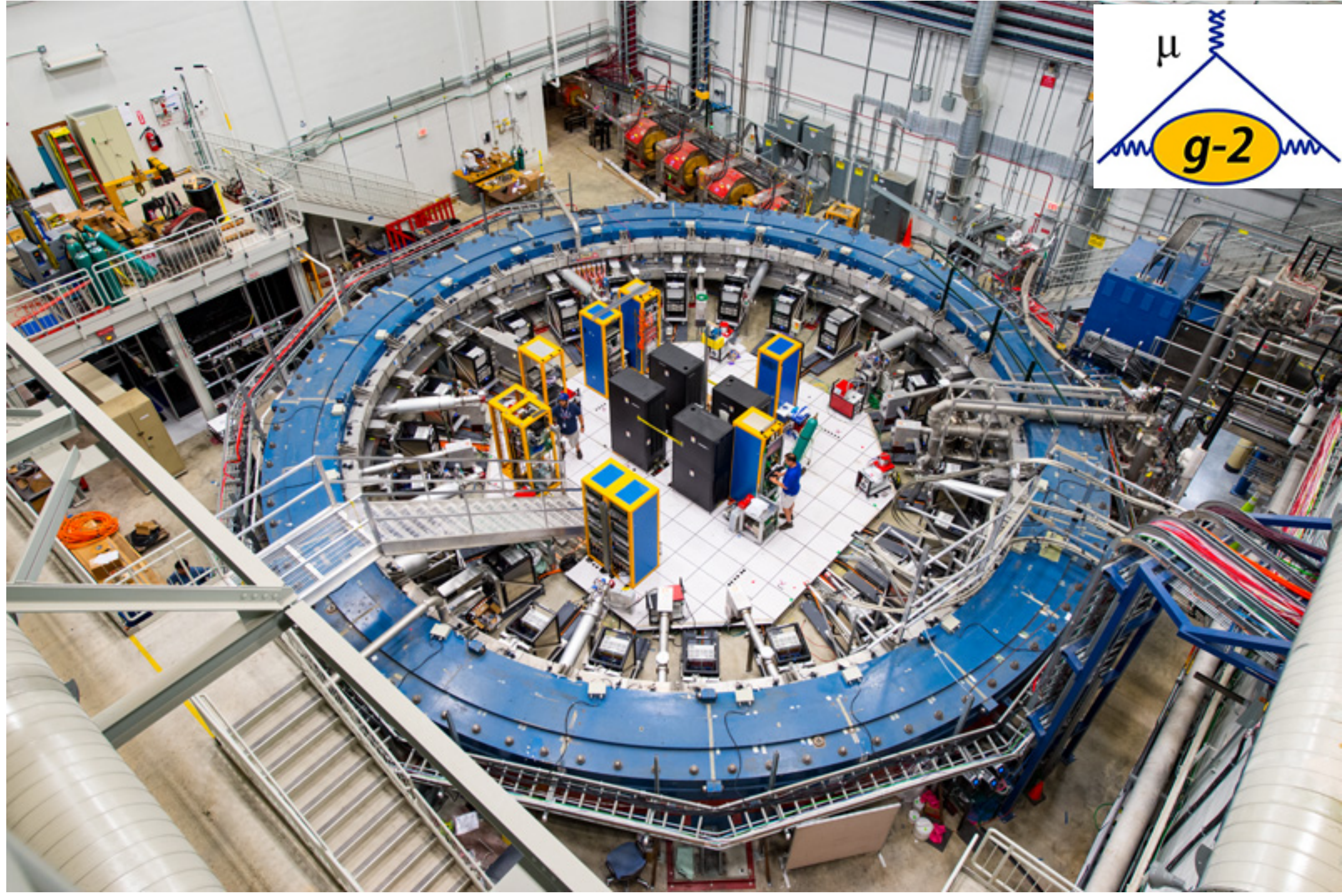


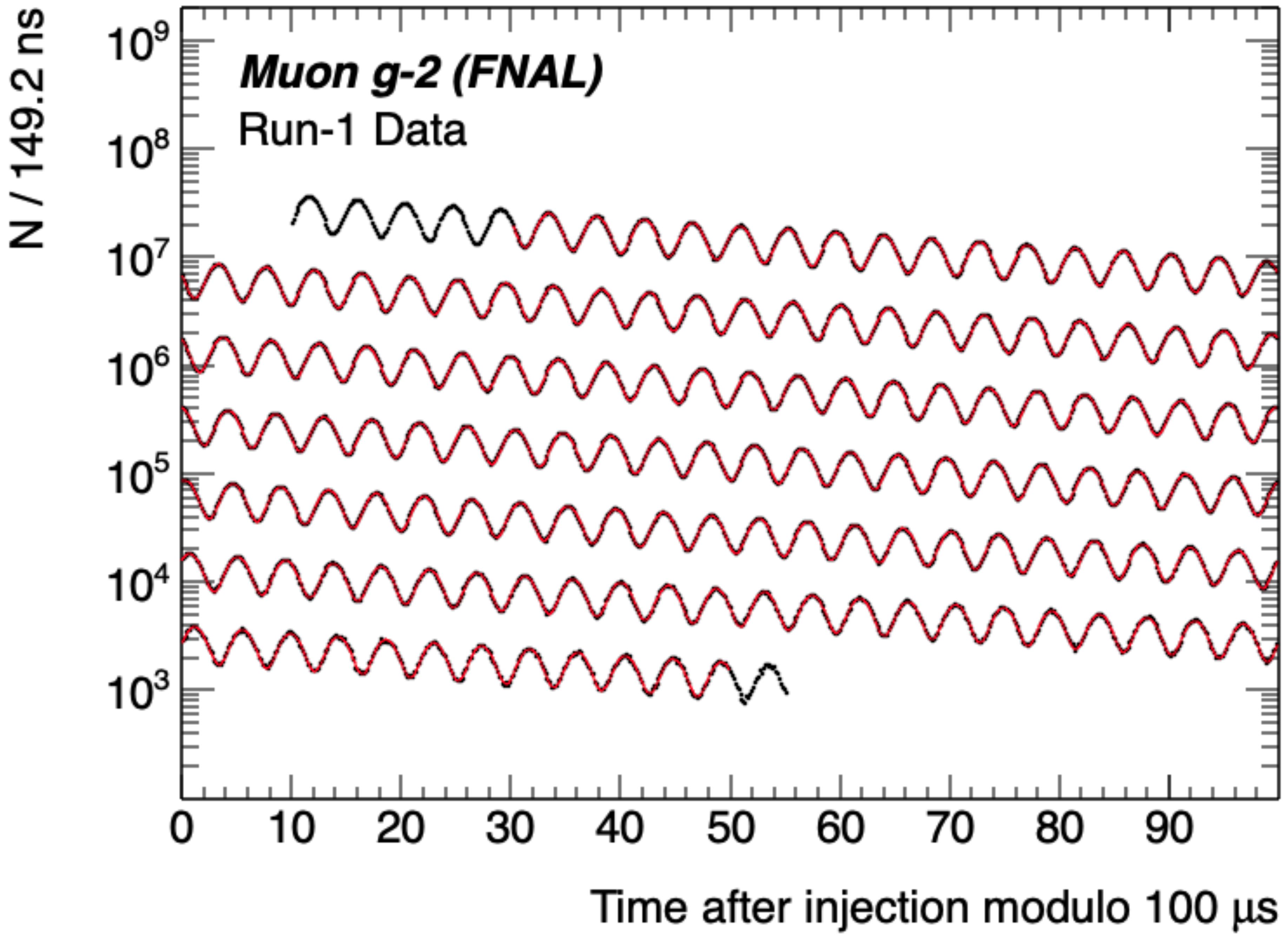
Most accurately verified prediction in the history of physics

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









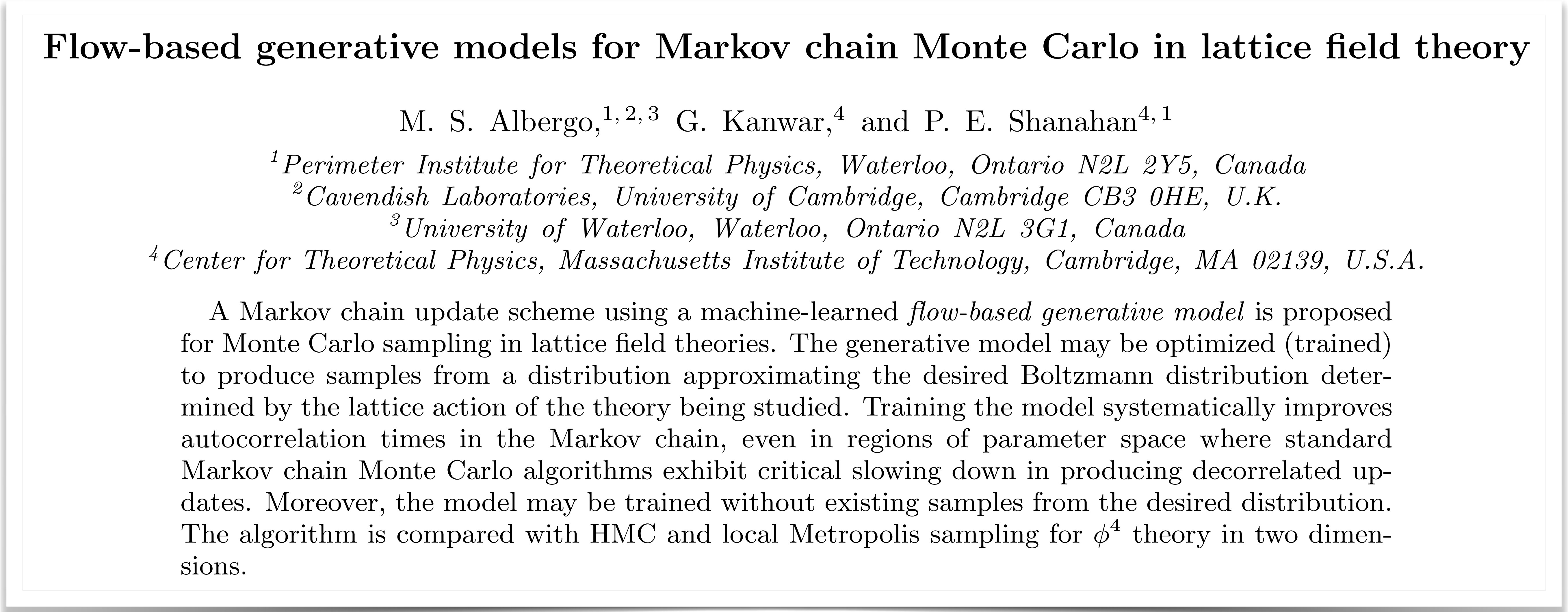
Flows for LQCD

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



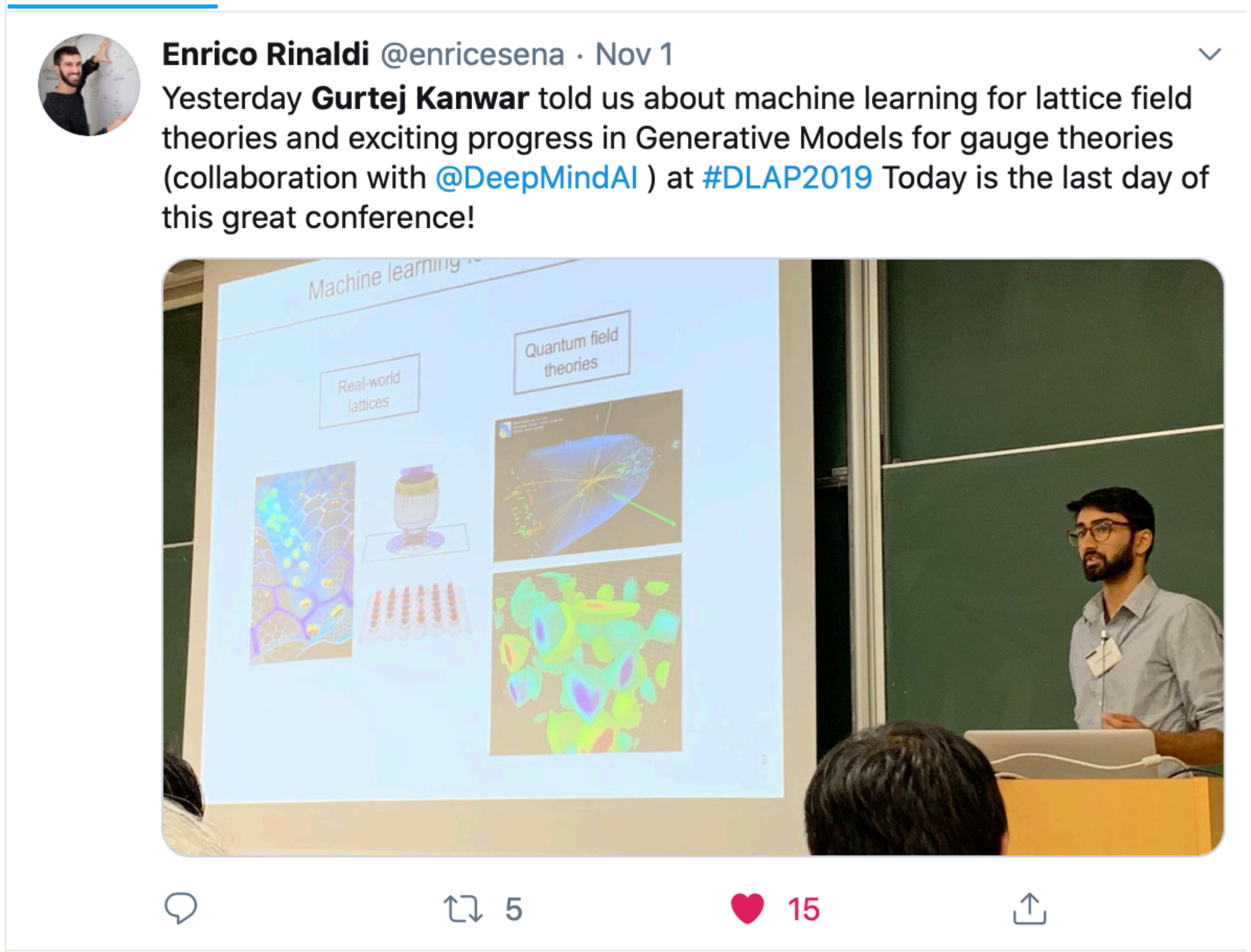
Flow-based generative models for Markov chain Monte Carlo in lattice field theory

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.

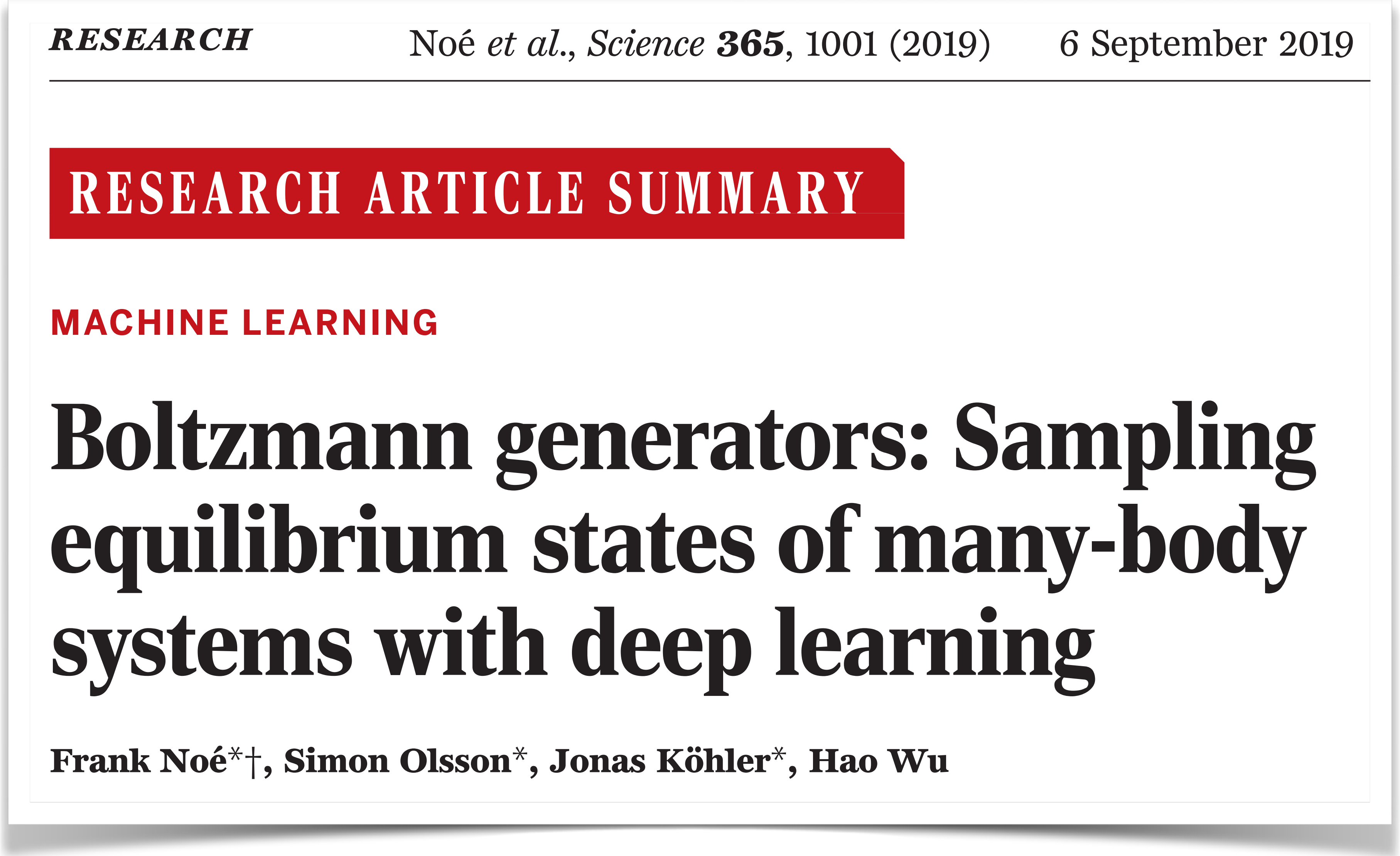
Hire Michael ↑



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!

Flows for molecular dynamics



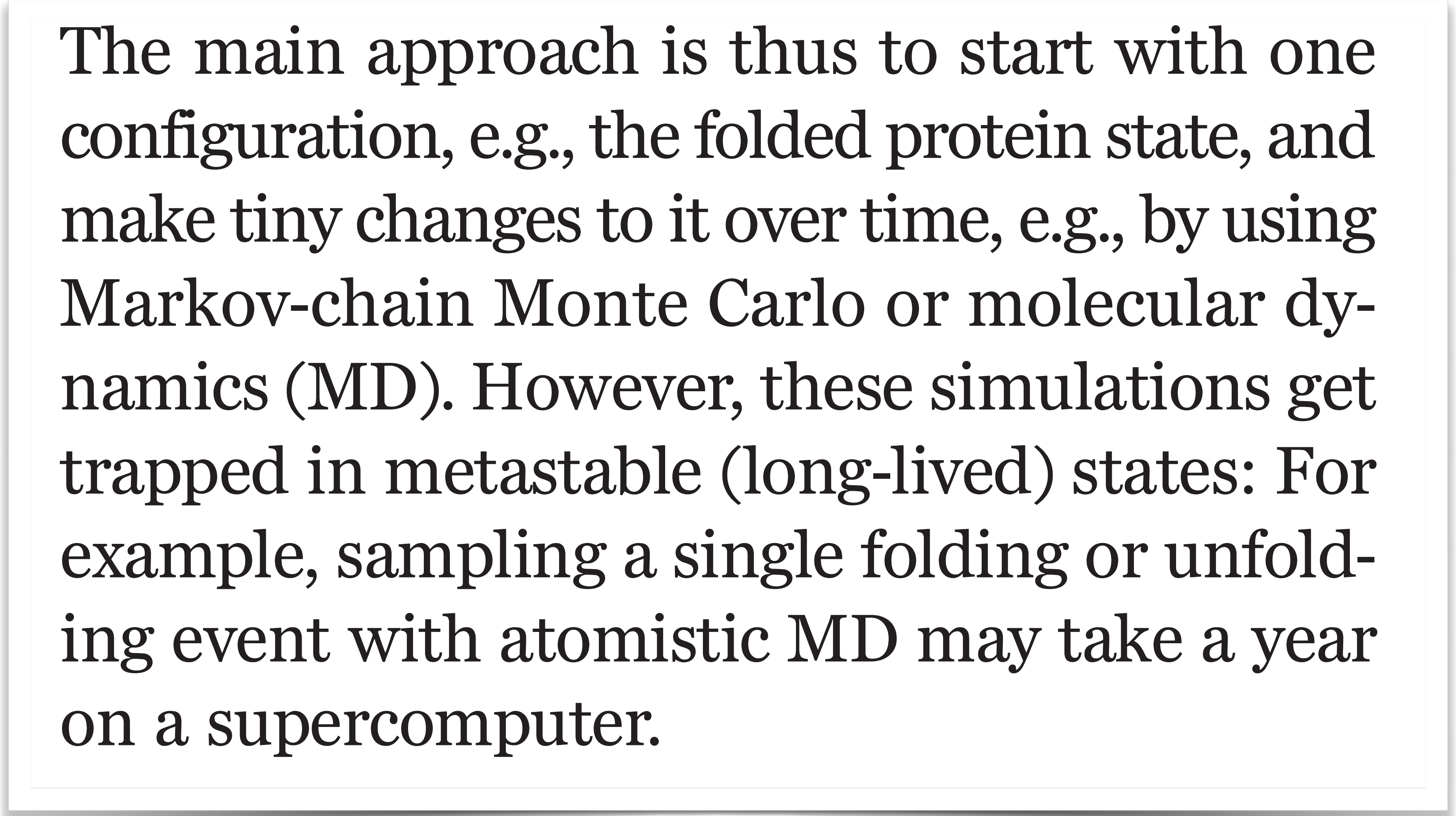
RESEARCH Noé *et al.*, *Science* **365**, 1001 (2019) 6 September 2019

RESEARCH ARTICLE SUMMARY

MACHINE LEARNING

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

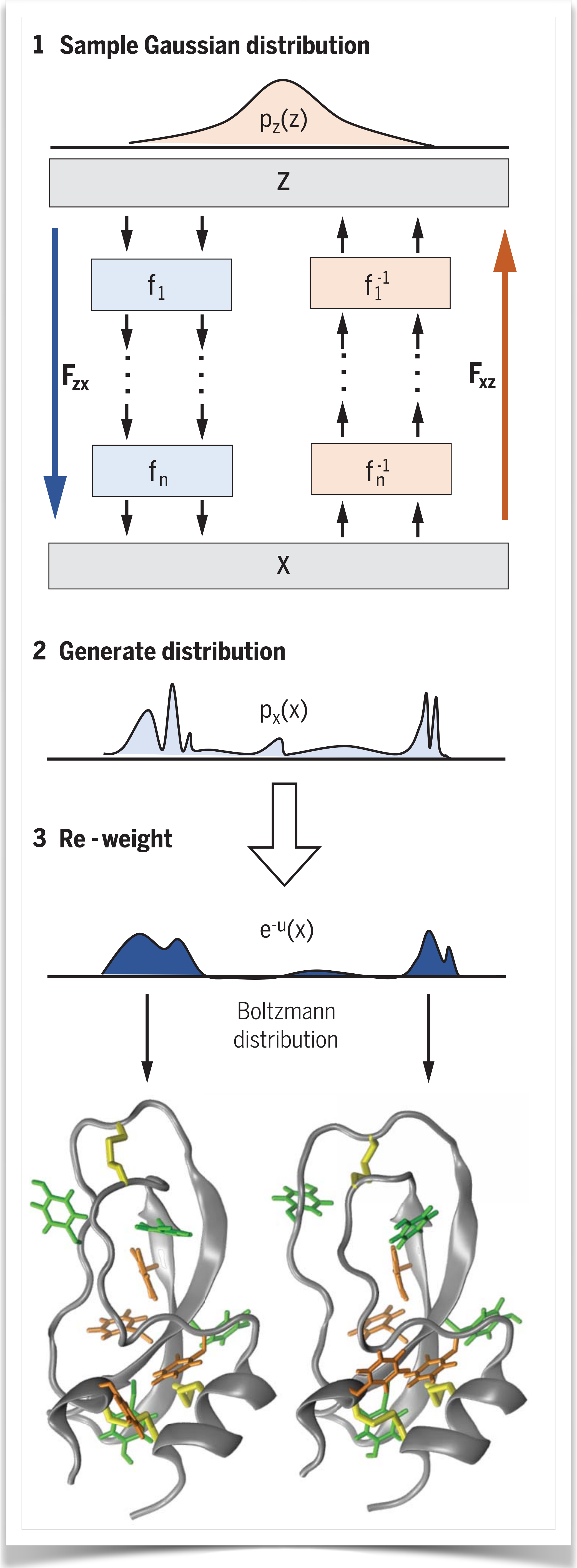
Frank Noé*†, Simon Olsson*, Jonas Köhler*, Hao Wu



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.

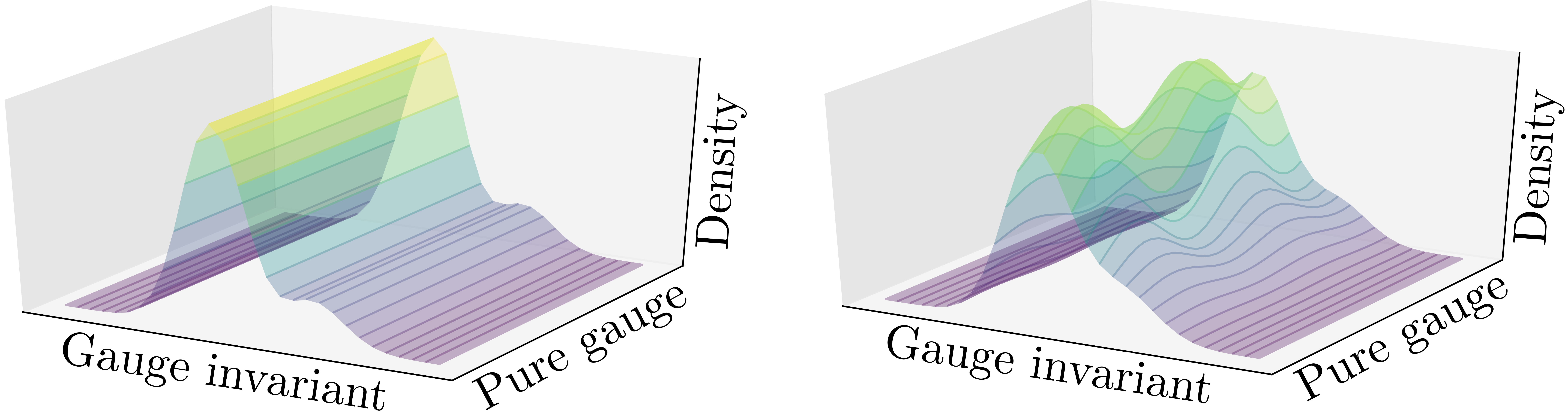


Space-time & Local Gauge Symmetry

The action is invariant to **local gauge transformations**, so the distribution is constant in those directions. **It's a huge product group!** $\mathrm{SU}(3)^{(\#\text{lattice sites})}$

Many more pure gauge degrees of freedom than physical ones

We would like to enforce this symmetry in the network, and not have to learn it.



Step 1: Flows on Spheres and Tori

[arXiv:2002.02428, ICML2020]

We designed flows on compact manifolds like Spheres and Tori that correspond to Lie groups:

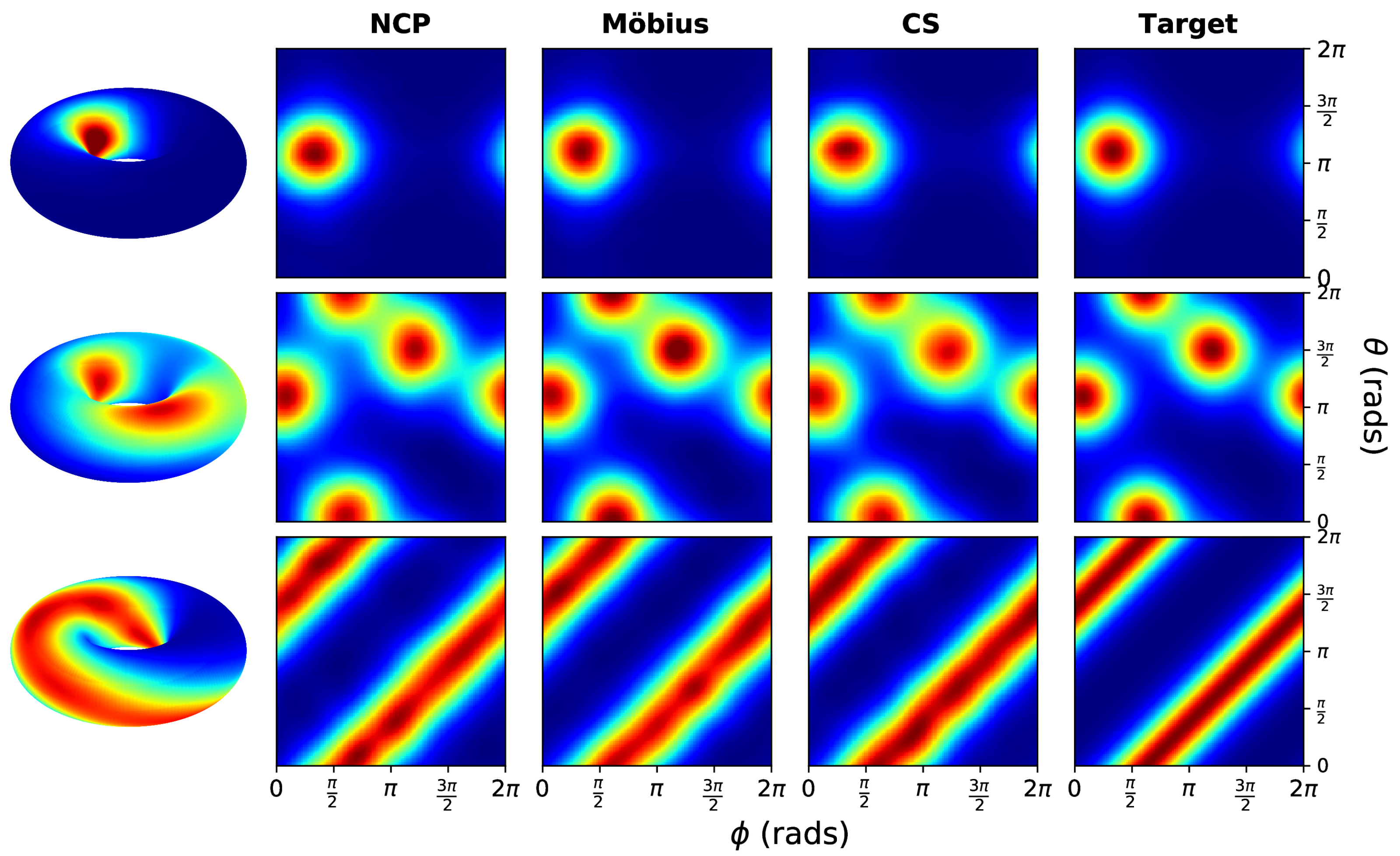
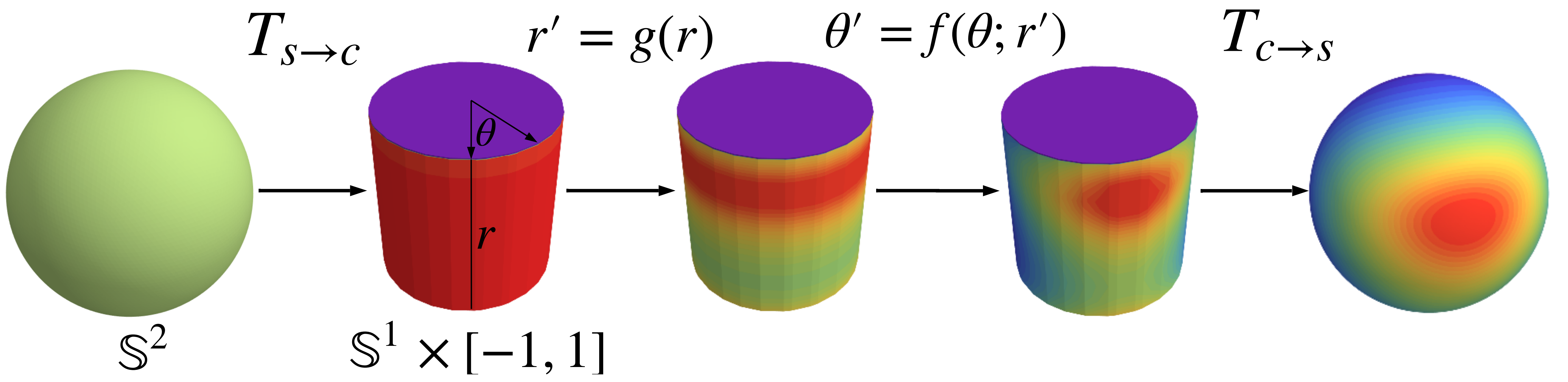
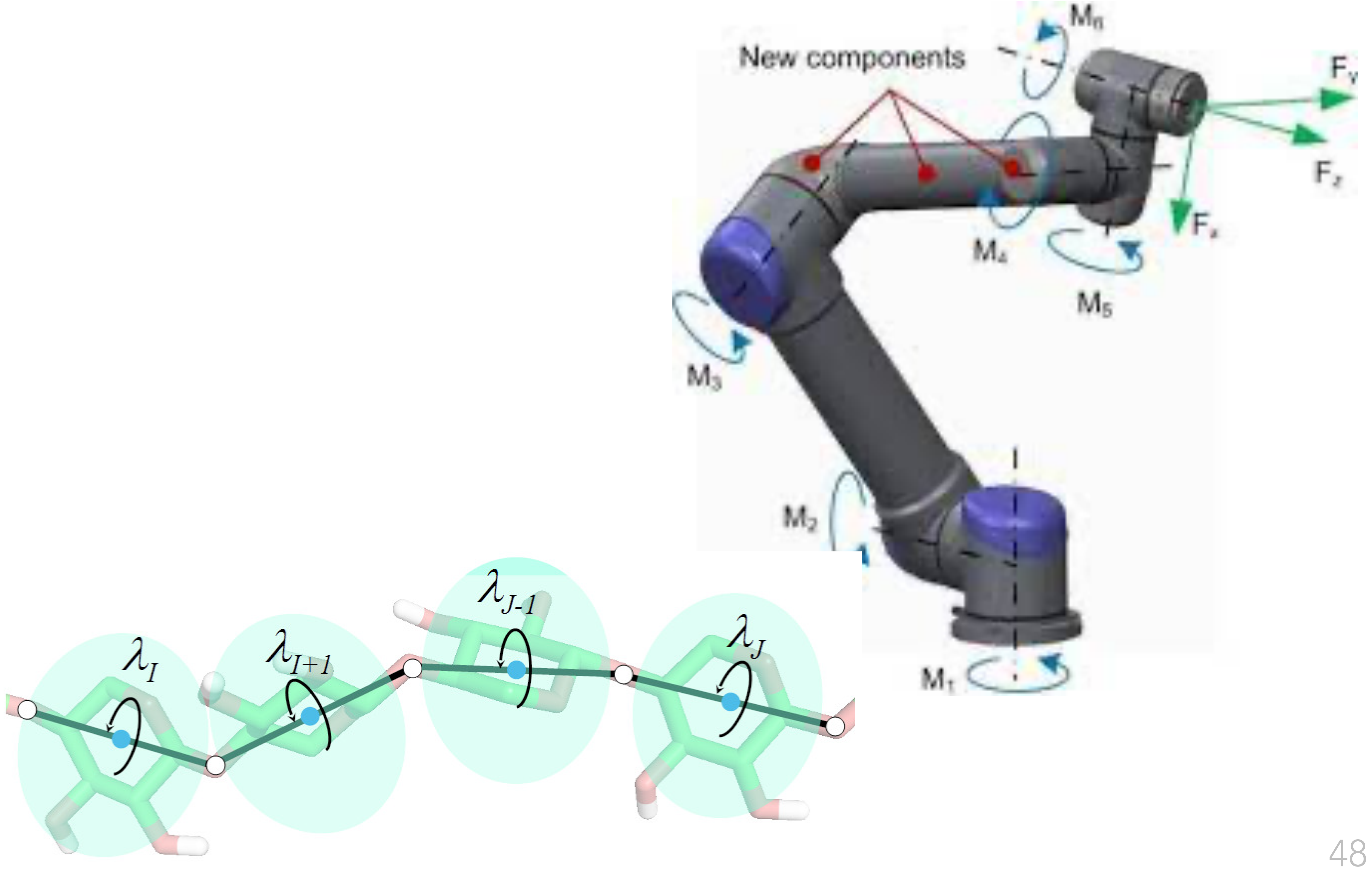


Figure 3. Learned densities on $\mathbb{T}^2$ using NCP, Möbius and CS flows. Densities shown on the torus are from NCP.





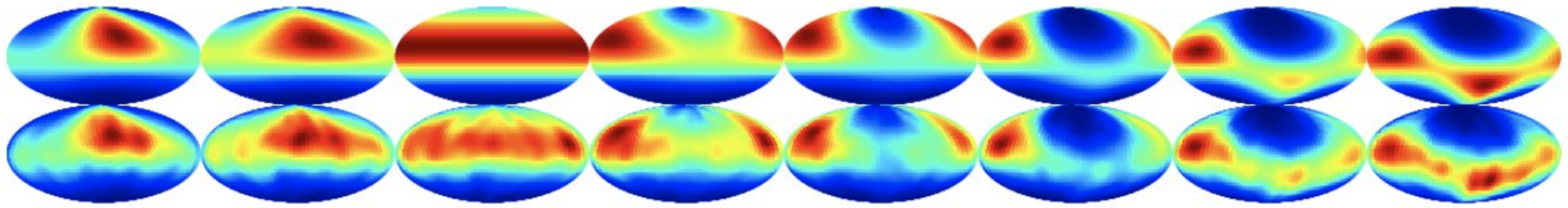
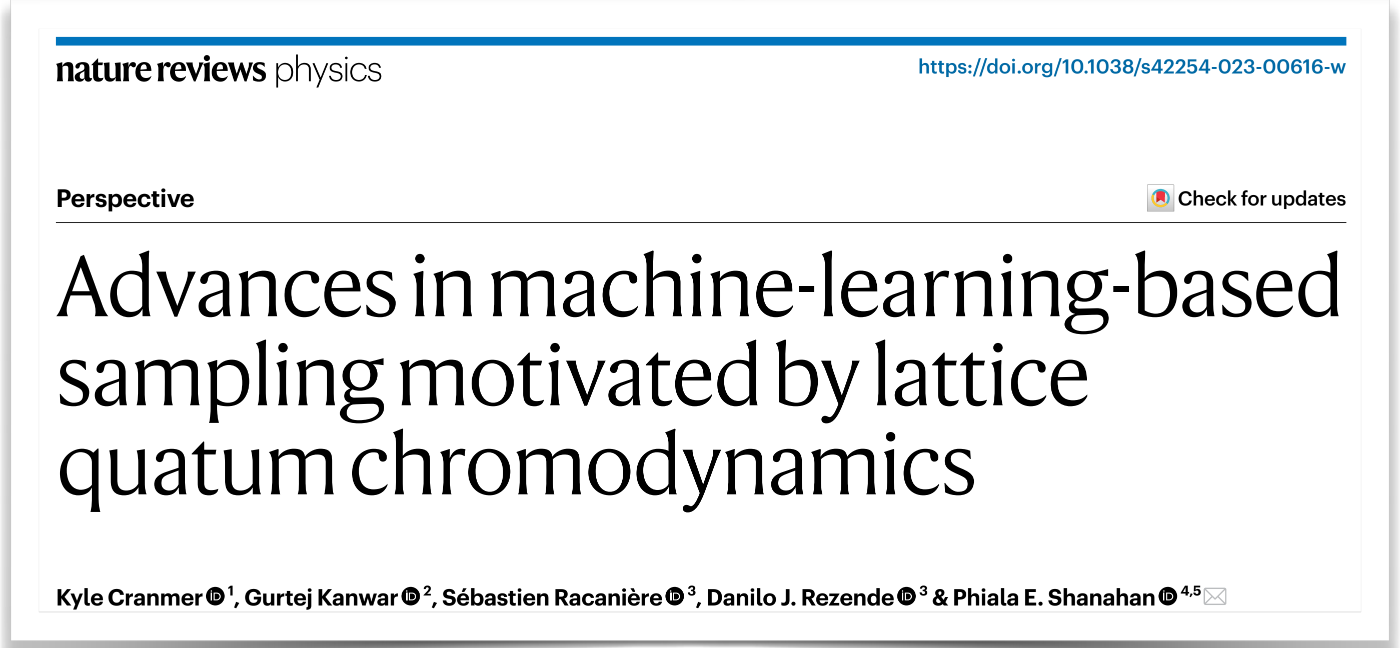


Figure 5. Learned multi-modal density on $SU(2) \equiv \mathbb{S}^3$ using the recursive flow. Each column shows an $\mathbb{S}^2$ slice of the $\mathbb{S}^3$ density

AI for Lattice Field Theory

Review ⇒



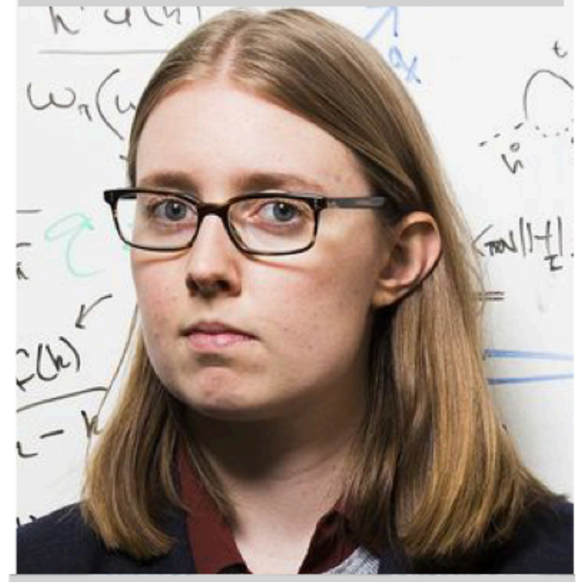
nature reviews physics https://doi.org/10.1038/s42254-023-00616-w

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]

Shanahan



Massachusetts Institute of Technology

Abbott

Massachusetts Institute of Technology

Hackett

Massachusetts Institute of Technology

Romero-López

Massachusetts Institute of Technology

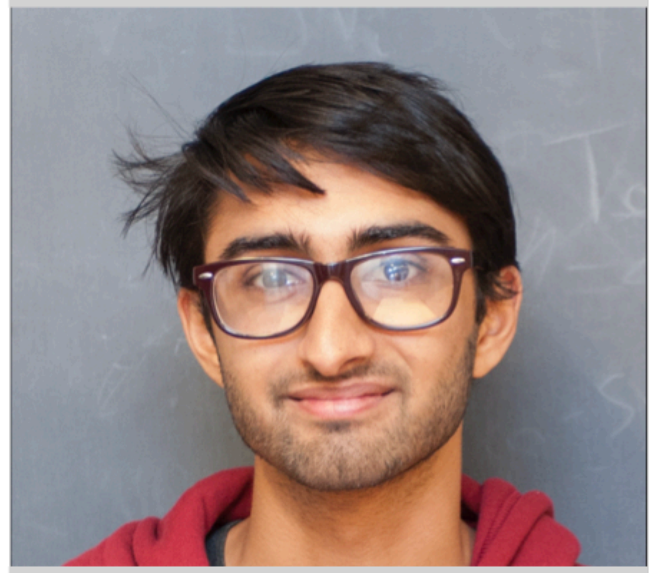
Boyda

Argonne NATIONAL LABORATORY → MIT

Urban

UNIVERSITÄT HEIDELBERG → MIT

Kanwar



MIT → u^b BERN

Botev

DeepMind

Matthews

DeepMind

Rezende

DeepMind

Racanière

DeepMind

Razavi

DeepMind

Albergo

NYU

Cranmer

WISCONSIN UNIVERSITY OF WISCONSIN–MADISON

Resource of Inspiration

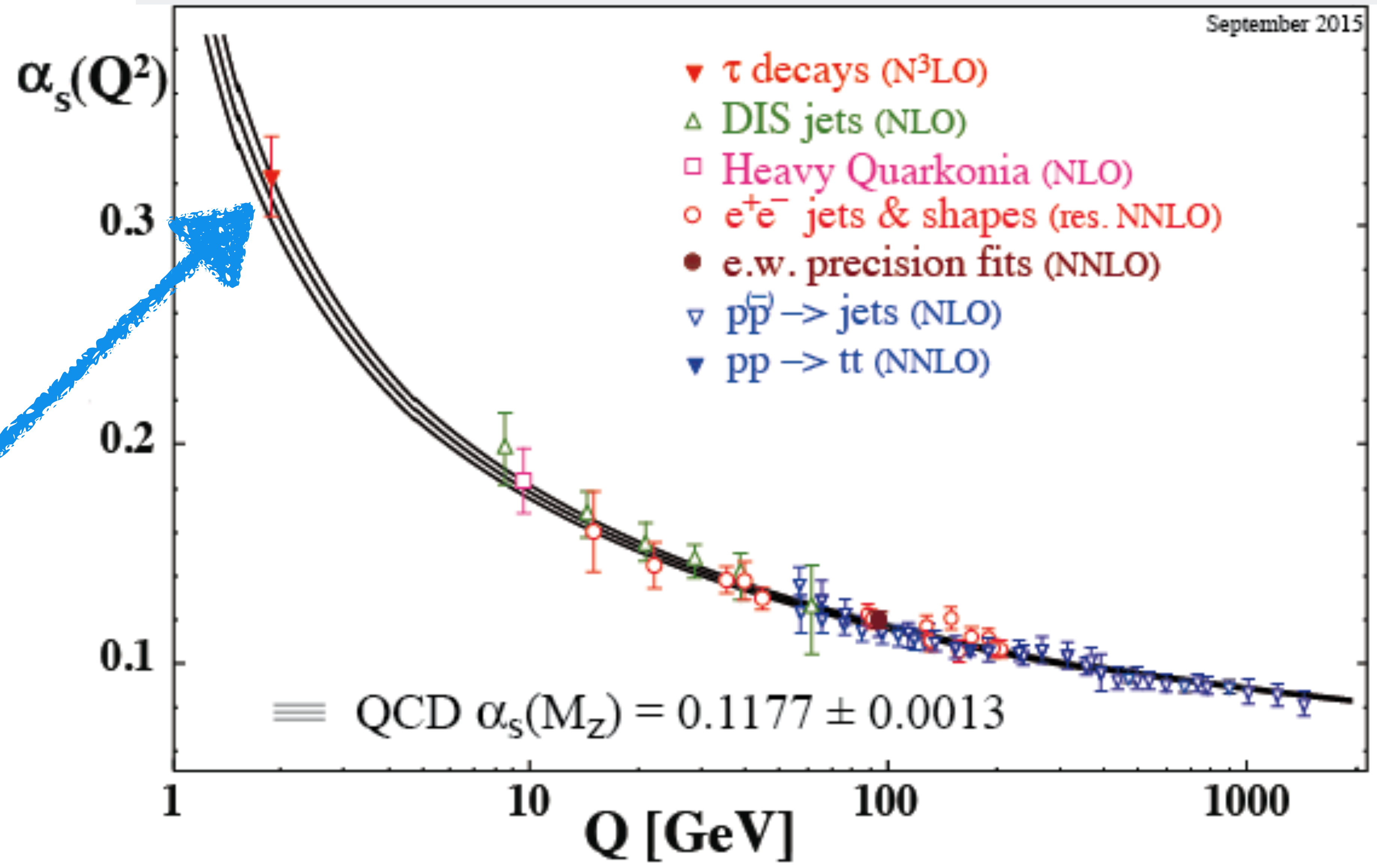
The strong force: Quantum Chromodynamics (QCD)

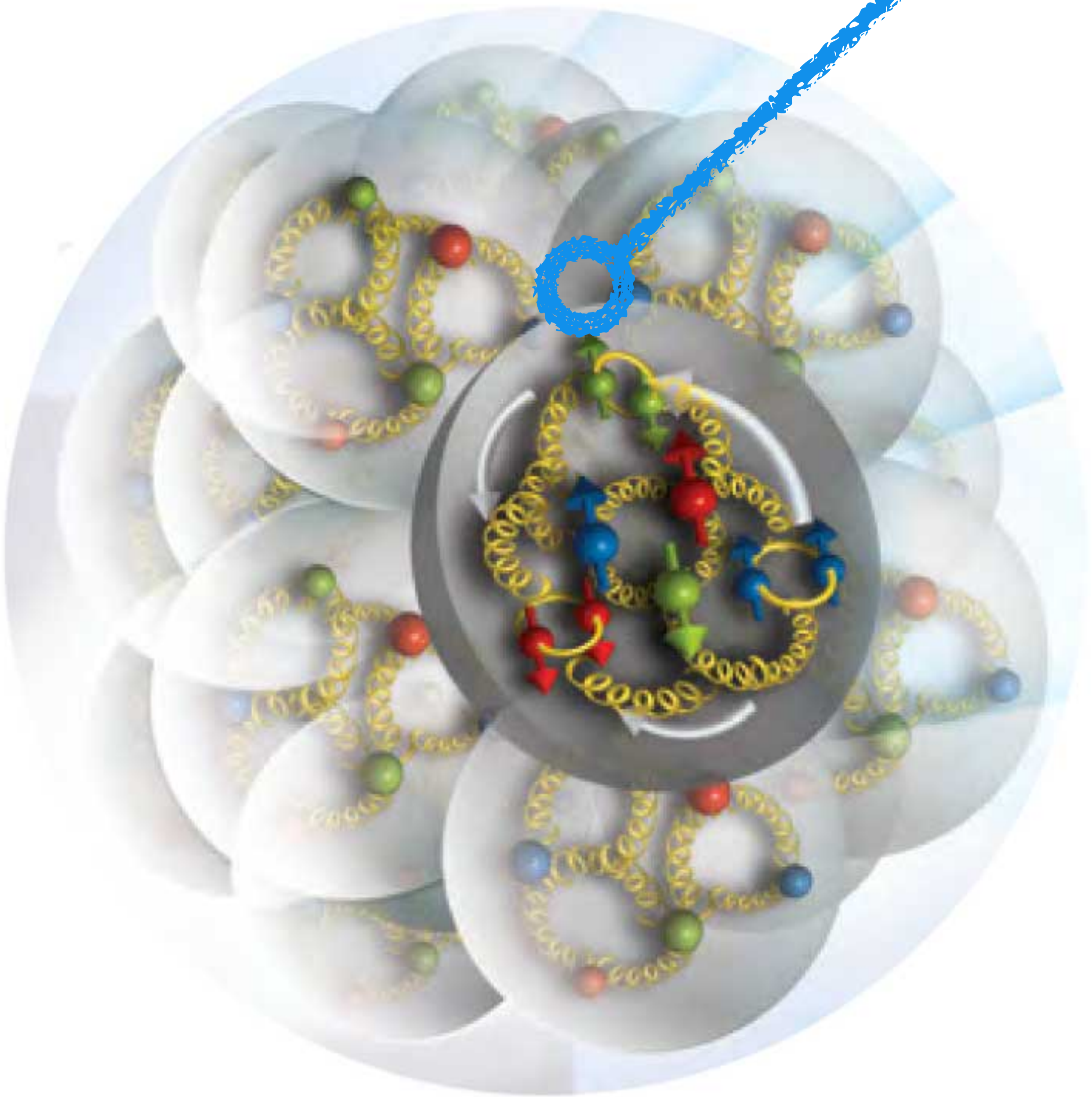
QCD is strong at at low-energies, no small coupling, perturbation theory fails.

Emergent phenomena: protons, pions, etc.

Interaction strength depends on energy

[Gross, Politzer, Wilczek, Nobel 2004]





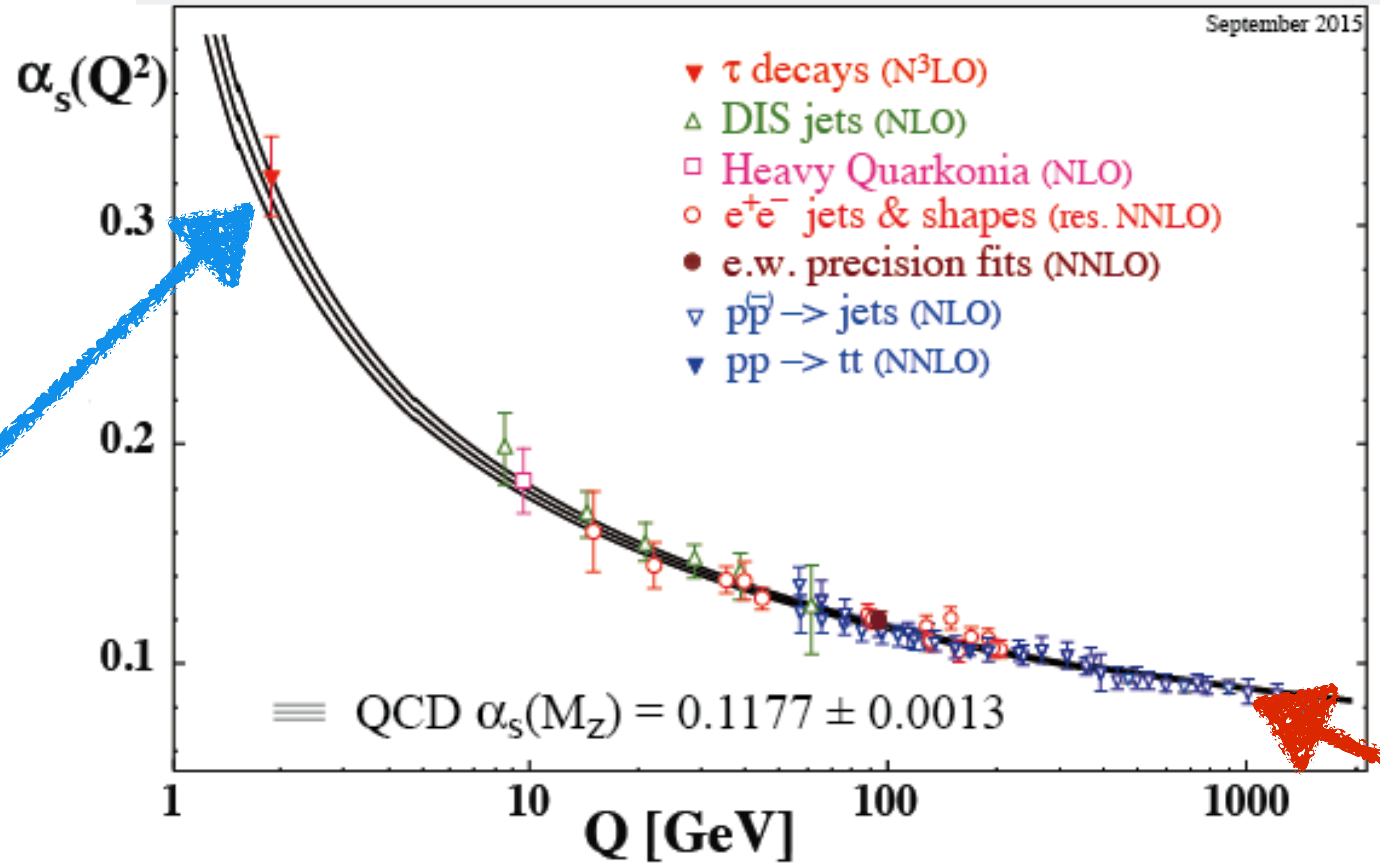
The strong force: Quantum Chromodynamics (QCD)

QCD is strong at at low-energies, no small coupling, perturbation theory fails.

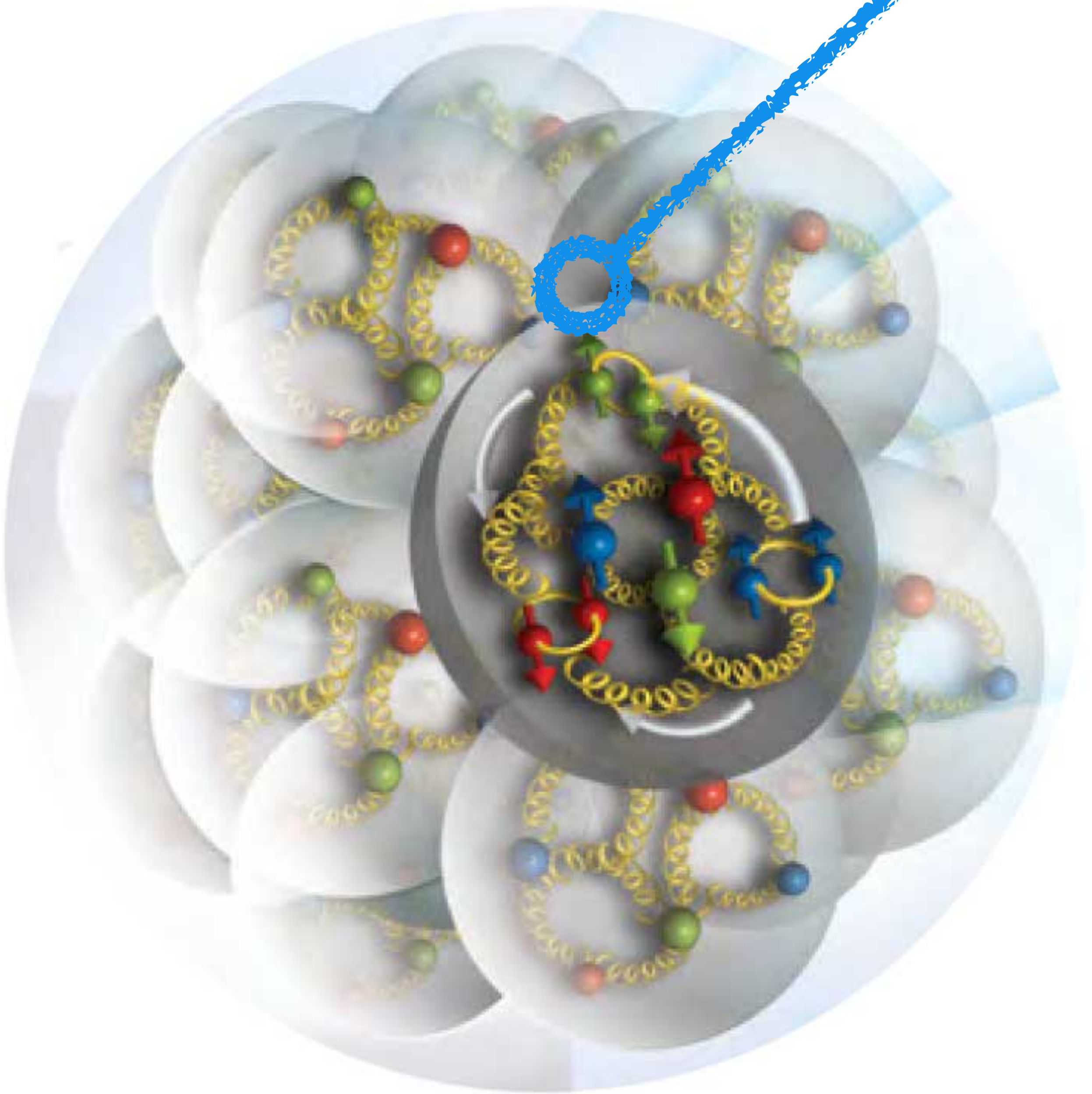
Emergent phenomena: protons, pions, etc.

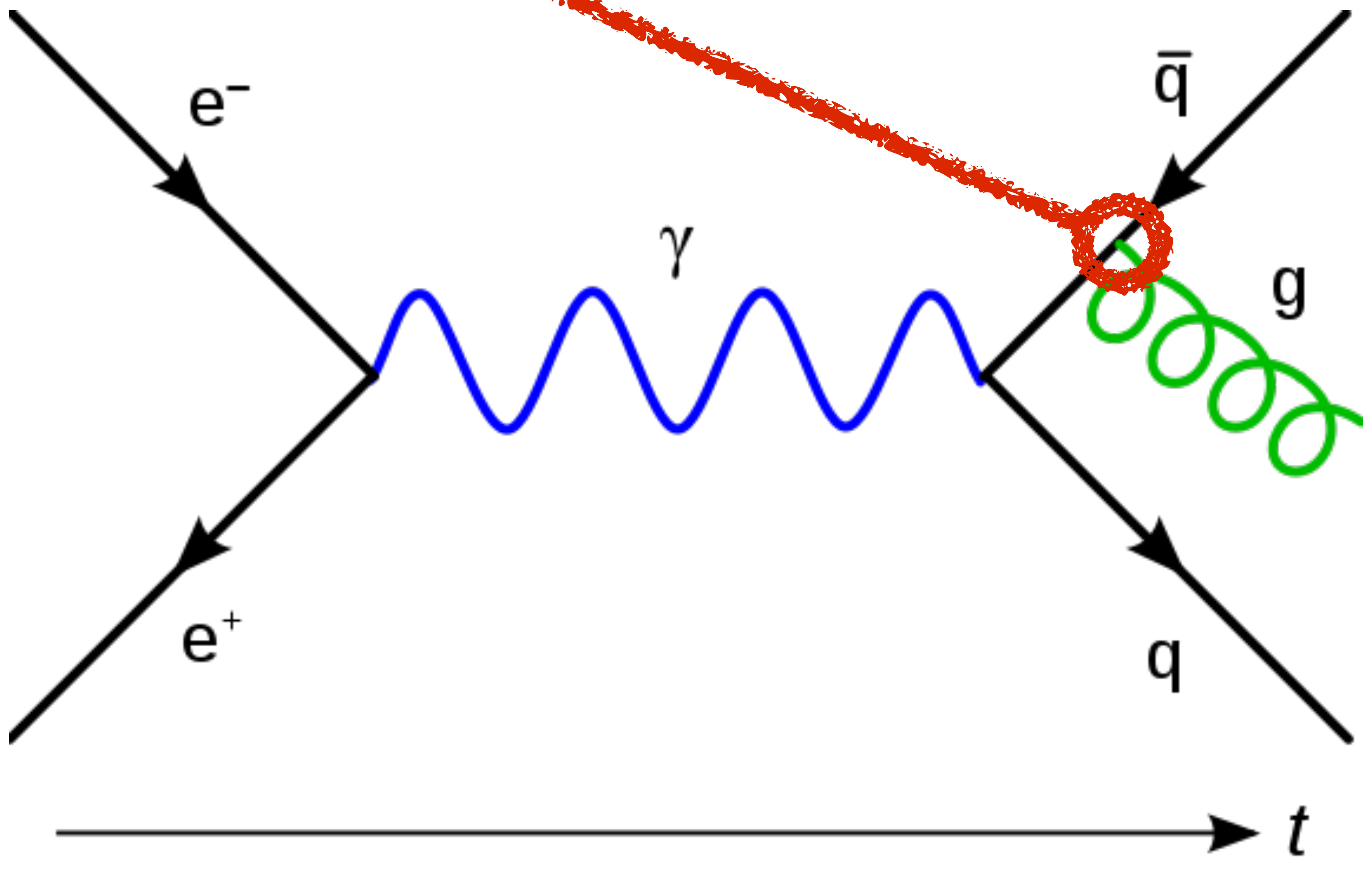
Interaction strength depends on energy

[Gross, Politzer, Wilczek, Nobel 2004]



QCD is weak at at high-energies, small coupling, perturbation theory works





> *Like the silicon chips of more recent years, the Feynman diagram was bringing computation to the masses.*

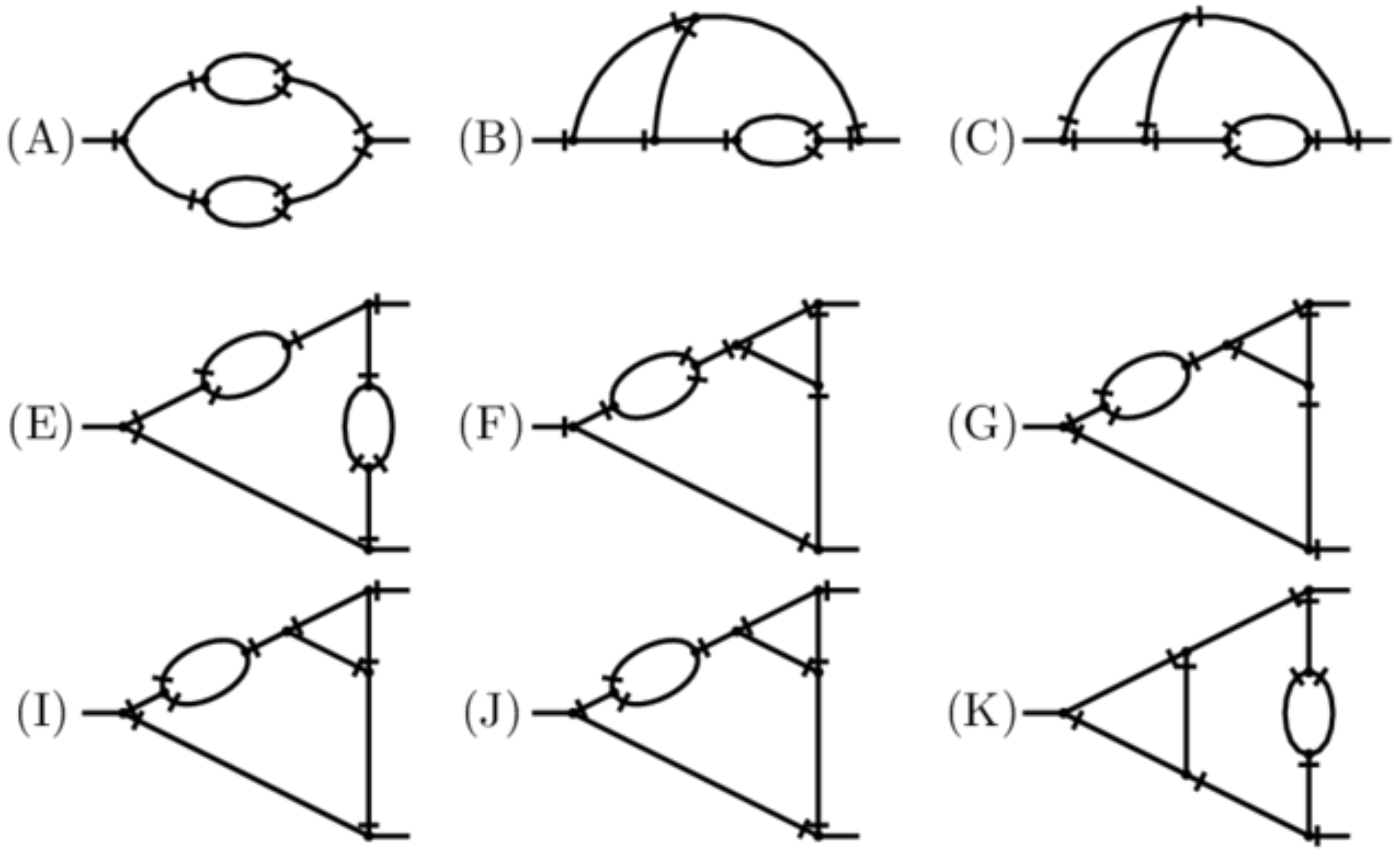
–JULIAN SCHWINGER

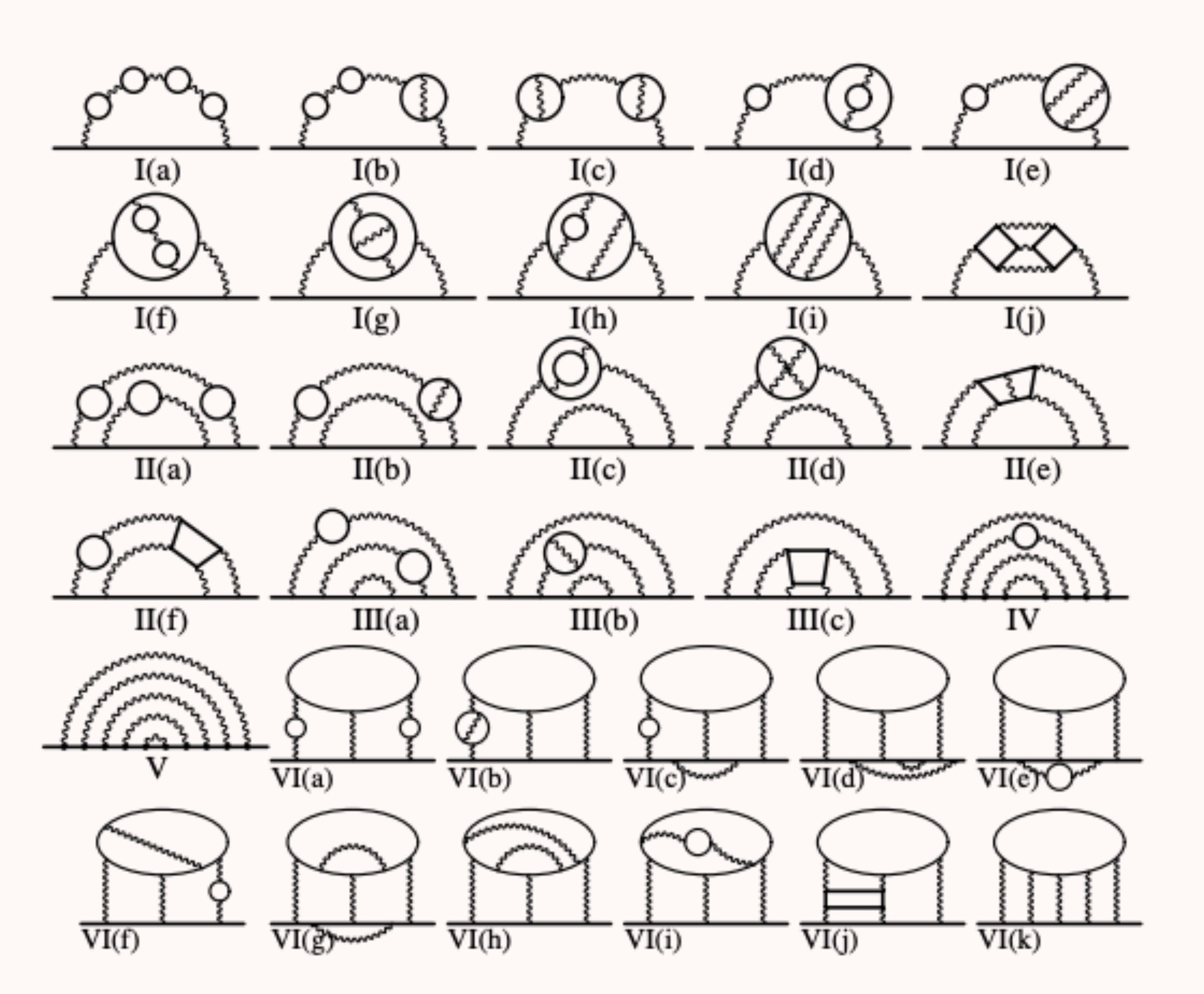
Feynman diagrams with loops

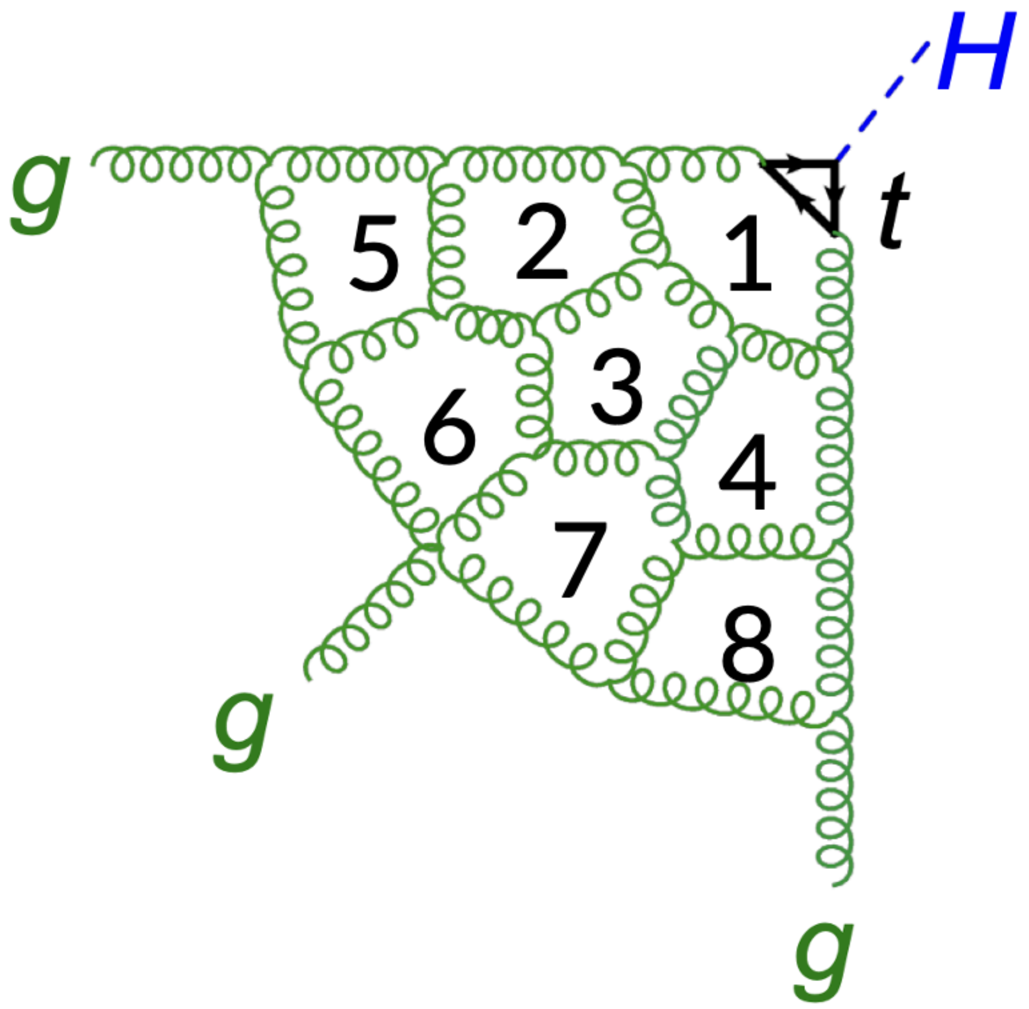
More precise calculations have more **loops**

- 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







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**

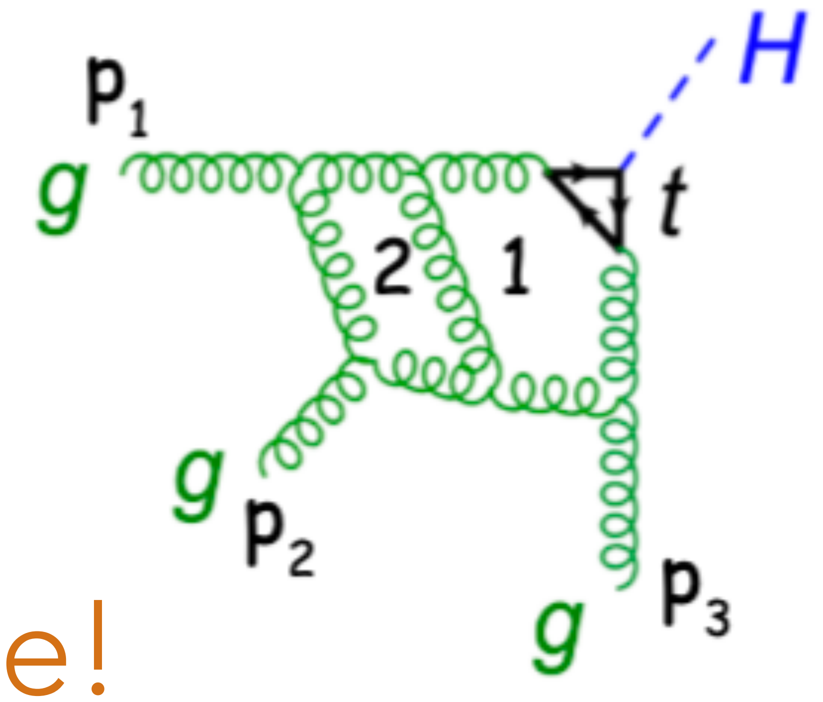
- The 6 letters $\{a, b, c, d, e, f\}$ encode the kinematics of the collision

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:



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

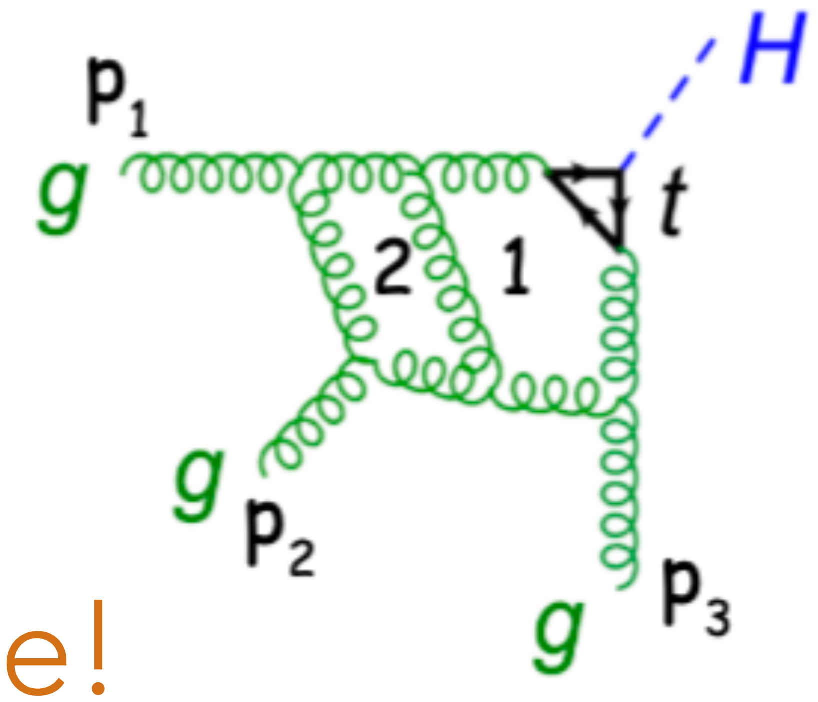
- Of the 6^4 =1296 possible terms, most are 0. Sparse, lots of structure!

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

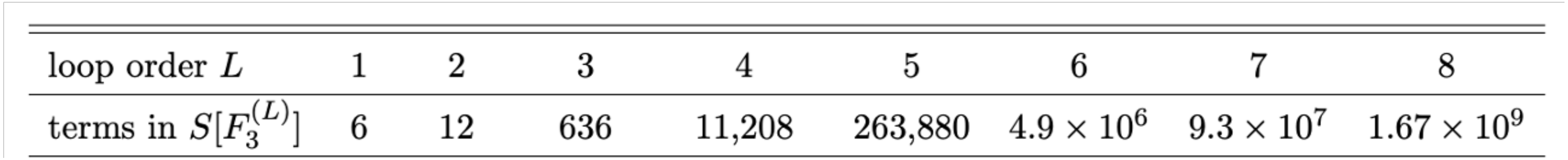
For example, in a particular theory called $\mathcal{N}$=4 Super Yang-Mills theory, the answer at 2-loops for a particular interaction is:



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

- 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

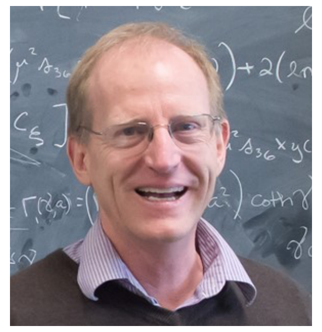


loop order L	1	2	3	4	5	6	7	8
terms in $S[F_3^{(L)}]$	6	12	636	11,208	263,880	4.9×10^6	9.3×10^7	1.67×10^9

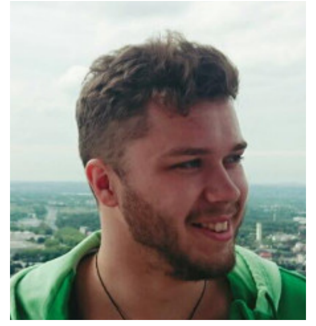
Can AI help?


Garrett Merz


Tianji Cai


Lance Dixon


Matthias Wilhelm


Niklas Nolte

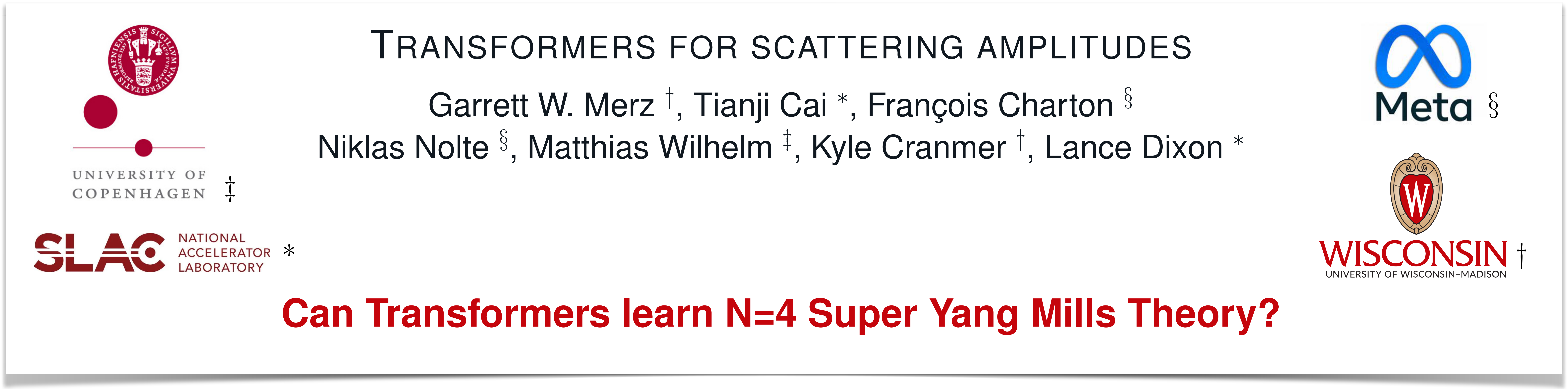

Francois Charton

Similar to an NP-complete problem:

- **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



TRANSFORMERS FOR SCATTERING AMPLITUDES

Garrett W. Merz [†], Tianji Cai [*], François Charton [§]
Niklas Nolte [§], Matthias Wilhelm [‡], Kyle Cranmer [†], Lance Dixon [*]

UNIVERSITY OF COPENHAGEN [‡]
SLAC NATIONAL ACCELERATOR LABORATORY [*]
Meta [§]
WISCONSIN UNIVERSITY OF WISCONSIN-MADISON [†]

Can Transformers learn N=4 Super Yang Mills Theory?

Paper & poster in ML and Physical Sciences workshop yesterday

Symbolic Regression

Complete the sequence:

- 1, 1, 2, 3, 5, 8, 13, 21, …

No i.i.d. data, only one example.

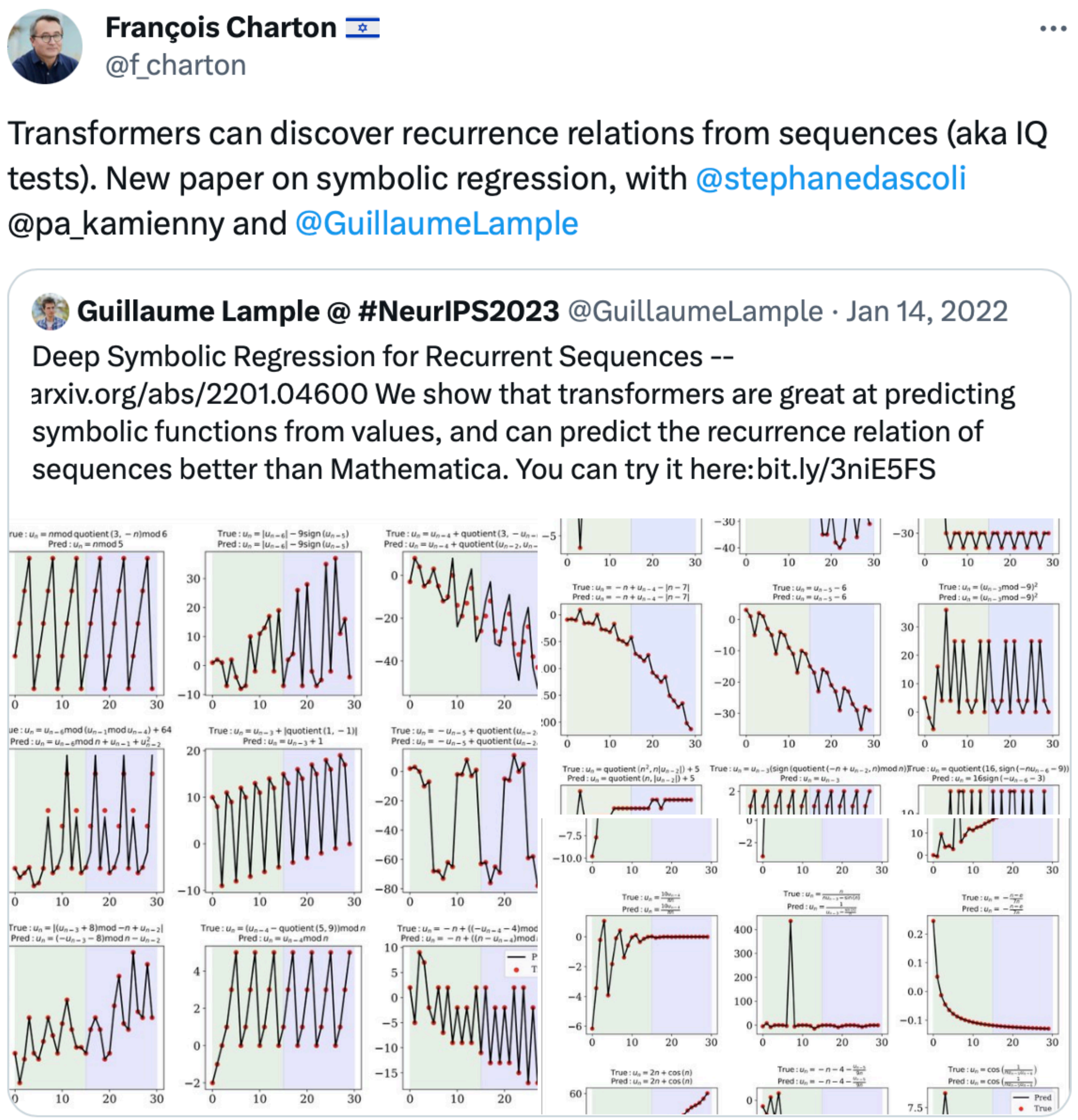
What is the recursion rule?

- $x_n = x_{n-1} + x_{n-2}$

Our problem is much harder:

- each element of the sequence is a "tensor" growing in dimension

$$\mathbb{Z}^{6^2} \to \mathbb{Z}^{6^4} \to \mathbb{Z}^{6^6} \to \mathbb{Z}^{6^8} \to \ldots \mathbb{Z}^{6^{2L}}$$



François Charton
@f_charton

Transformers can discover recurrence relations from sequences (aka IQ tests). New paper on symbolic regression, with @stephanedascoli @pa_kamienny and @GuillaumeLample

Guillaume Lample @ #NeurIPS2023 @GuillaumeLample · Jan 14, 2022

Deep Symbolic Regression for Recurrent Sequences -- arxiv.org/abs/2201.04600 We show that transformers are great at predicting symbolic functions from values, and can predict the recurrence relation of sequences better than Mathematica. You can try it here:bit.ly/3niE5FS

Building Community

The ML4Jets Workshops

2017

2020

2023

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

Hammers & Nails

NeurIPS workshops

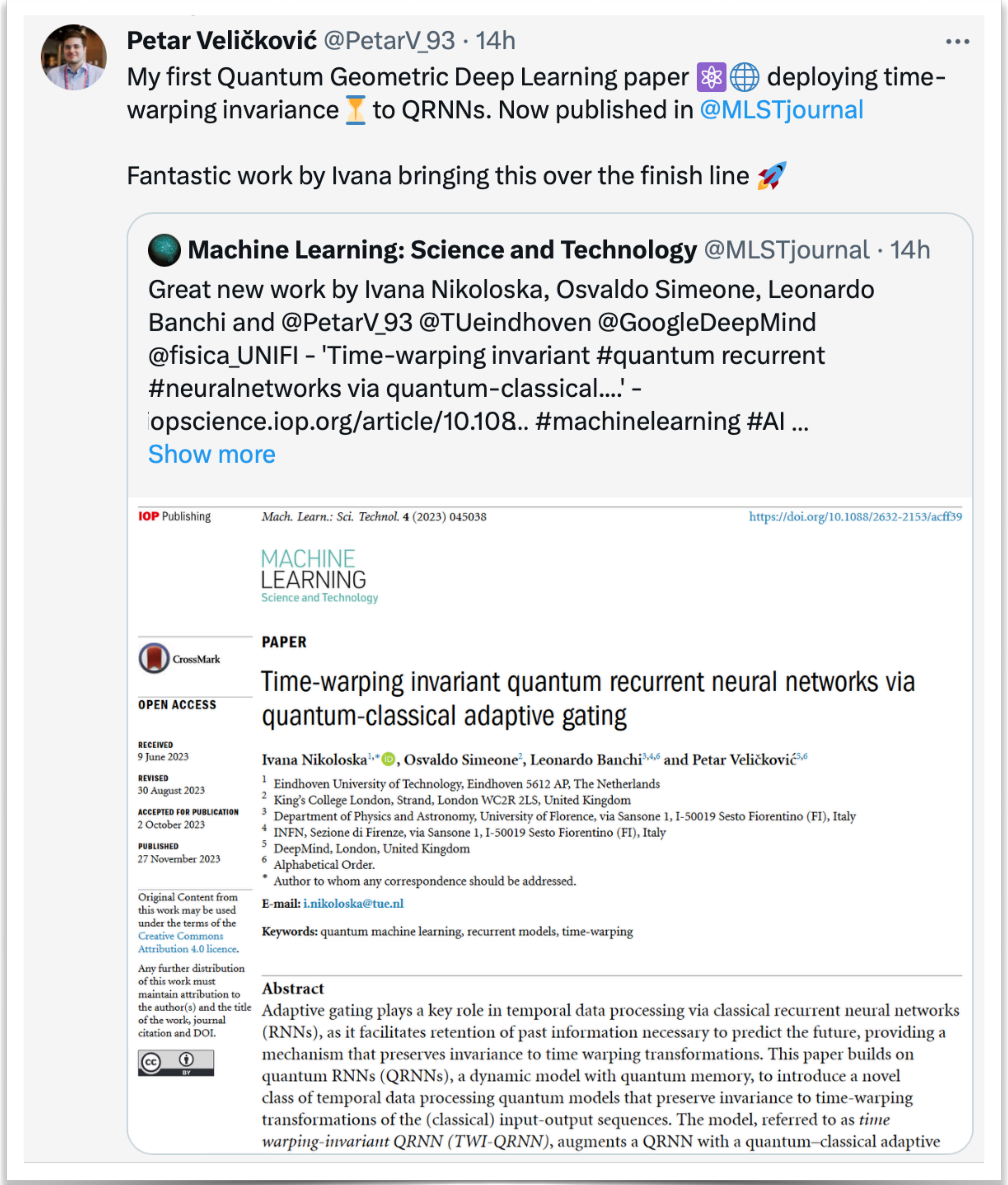
Machine Learning and the Physical Sciences

Workshop at the 37th conference on Neural Information Processing Systems (NeurIPS)

December 15, 2023

A journal to consider

Machine Learning: Science and Technology is a multidisciplinary open access journal that bridges the application of machine learning across the sciences with advances in machine learning methods and theory as motivated by physical insights. Disclosure: I'm the current Editor-in-Chief



Petar Veličković @PetarV_93 · 14h

My first Quantum Geometric Deep Learning paper ⚛🌐 deploying time-warping invariance ⏳ to QRNNs. Now published in @MLSTjournal

Fantastic work by Ivana bringing this over the finish line 🚀

Machine Learning: Science and Technology @MLSTjournal · 14h

Great new work by Ivana Nikoloska, Osvaldo Simeone, Leonardo Banchi and @PetarV_93 @TUeindhoven @GoogleDeepMind @fisica_UNIFI - 'Time-warping invariant #quantum recurrent #neuralnetworks via quantum-classical....' - iopscience.iop.org/article/10.108.. #machinelearning #AI ...
Show more

IOP Publishing *Mach. Learn.: Sci. Technol.* **4** (2023) 045038 https://doi.org/10.1088/2632-2153/acff39

MACHINE
LEARNING
Science and Technology

CrossMark

OPEN ACCESS

RECEIVED
9 June 2023

REVISED
30 August 2023

ACCEPTED FOR PUBLICATION
2 October 2023

PUBLISHED
27 November 2023

Original Content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence.

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

PAPER

Time-warping invariant quantum recurrent neural networks via quantum-classical adaptive gating

Ivana Nikoloska[1,*], Osvaldo Simeone[2], Leonardo Banchi[3,4,6] and Petar Veličković[5,6]

[1] Eindhoven University of Technology, Eindhoven 5612 AP, The Netherlands
[2] King's College London, Strand, London WC2R 2LS, United Kingdom
[3] Department of Physics and Astronomy, University of Florence, via Sansone 1, I-50019 Sesto Fiorentino (FI), Italy
[4] INFN, Sezione di Firenze, via Sansone 1, I-50019 Sesto Fiorentino (FI), Italy
[5] DeepMind, London, United Kingdom
[6] Alphabetical Order.
[*] Author to whom any correspondence should be addressed.

E-mail: i.nikoloska@tue.nl

Keywords: quantum machine learning, recurrent models, time-warping

Abstract

Adaptive gating plays a key role in temporal data processing via classical recurrent neural networks (RNNs), as it facilitates retention of past information necessary to predict the future, providing a mechanism that preserves invariance to time warping transformations. This paper builds on quantum RNNs (QRNNs), a dynamic model with quantum memory, to introduce a novel class of temporal data processing quantum models that preserve invariance to time-warping transformations of the (classical) input-output sequences. The model, referred to as *time warping-invariant QRNN (TWI-QRNN)*, augments a QRNN with a quantum–classical adaptive

PAPER • OPEN ACCESS

Self-referencing embedded strings (SELFIES): A 100% robust molecular string representation

Mario Krenn[1,2,3], Florian Häse[1,2,3,4], AkshatKumar Nigam[2], Pascal Friederich[2,5] and Alan Aspuru-Guzik[1,2,3,6]

Published 28 October 2020 • © 2020 The Author(s). Published by IOP Publishing Ltd

Machine Learning: Science and Technology, Volume 1, Number 4

Citation Mario Krenn *et al* 2020 *Mach. Learn.: Sci. Technol.* **1** 045024

DOI 10.1088/2632-2153/aba947

PAPER • OPEN ACCESS

Data-driven discovery of Koopman eigenfunctions for control

Eurika Kaiser[3,1], J Nathan Kutz[2] and Steven L Brunton[1]

Published 14 June 2021 • © 2021 The Author(s). Published by IOP Publishing Ltd

Machine Learning: Science and Technology, Volume 2, Number 3

Citation Eurika Kaiser *et al* 2021 *Mach. Learn.: Sci. Technol.* **2** 035023

DOI 10.1088/2632-2153/abf0f5

PAPER • OPEN ACCESS

Graph machine learning framework for depicting wavefunction on interface

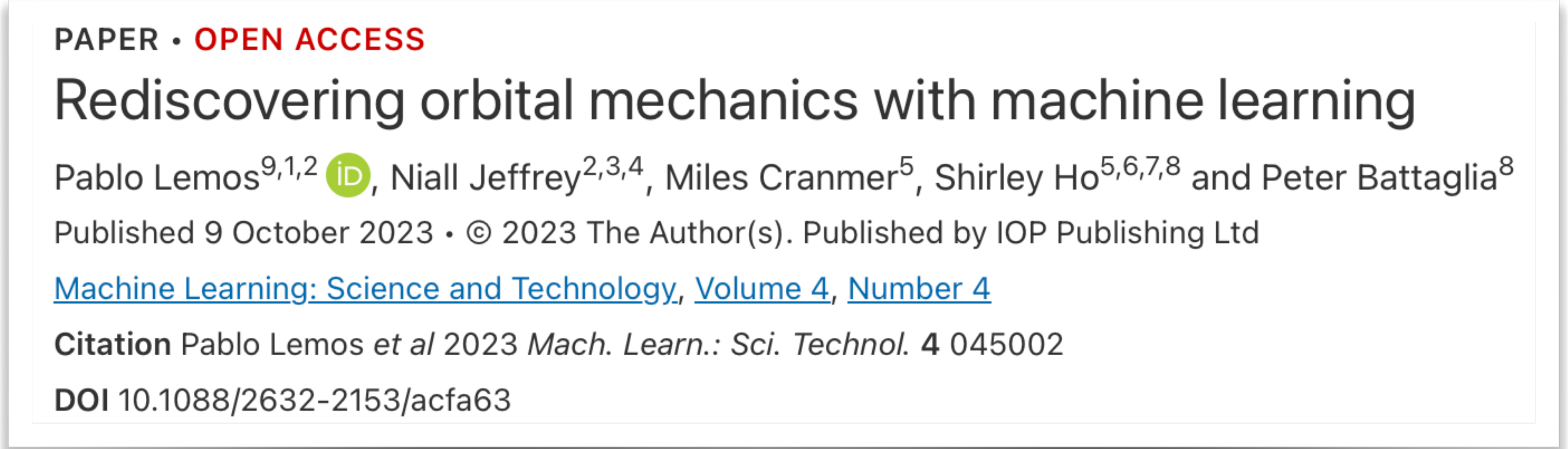
Ao Wu[1,2], Li Liu[1,2], Zifeng Wang[1,2], Shurong Pan[1,2], Jiangxue Huang[1,2], Qijun Huang[1,2], Jin He[1,2], Hao Wang[1,2] and Sheng Chang[3,1,2]

Published 27 November 2023 • © 2023 The Author(s). Published by IOP Publishing Ltd

Machine Learning: Science and Technology, Volume 4, Number 4

Citation Ao Wu *et al* 2023 *Mach. Learn.: Sci. Technol.* **4** 045037

DOI 10.1088/2632-2153/ad0937



PAPER • OPEN ACCESS

Rediscovering orbital mechanics with machine learning

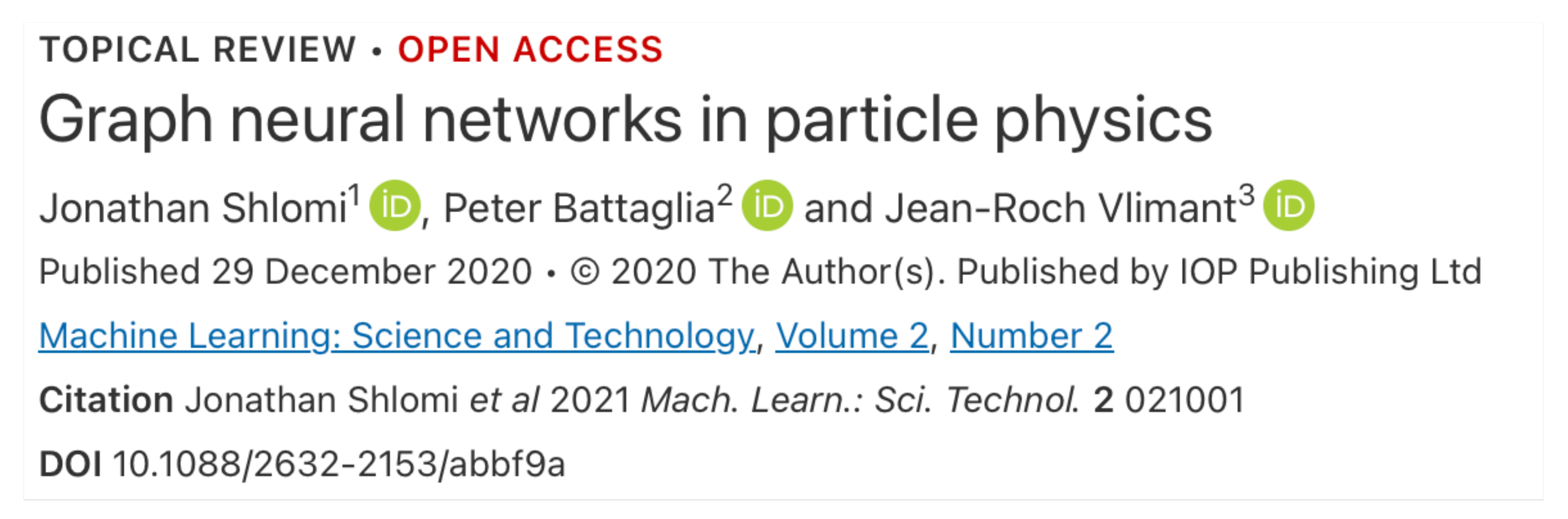
Pablo Lemos[9,1,2], Niall Jeffrey[2,3,4], Miles Cranmer[5], Shirley Ho[5,6,7,8] and Peter Battaglia[8]

Published 9 October 2023 • © 2023 The Author(s). Published by IOP Publishing Ltd

Machine Learning: Science and Technology, Volume 4, Number 4

Citation Pablo Lemos *et al* 2023 *Mach. Learn.: Sci. Technol.* **4** 045002

DOI 10.1088/2632-2153/acfa63



TOPICAL REVIEW • OPEN ACCESS

Graph neural networks in particle physics

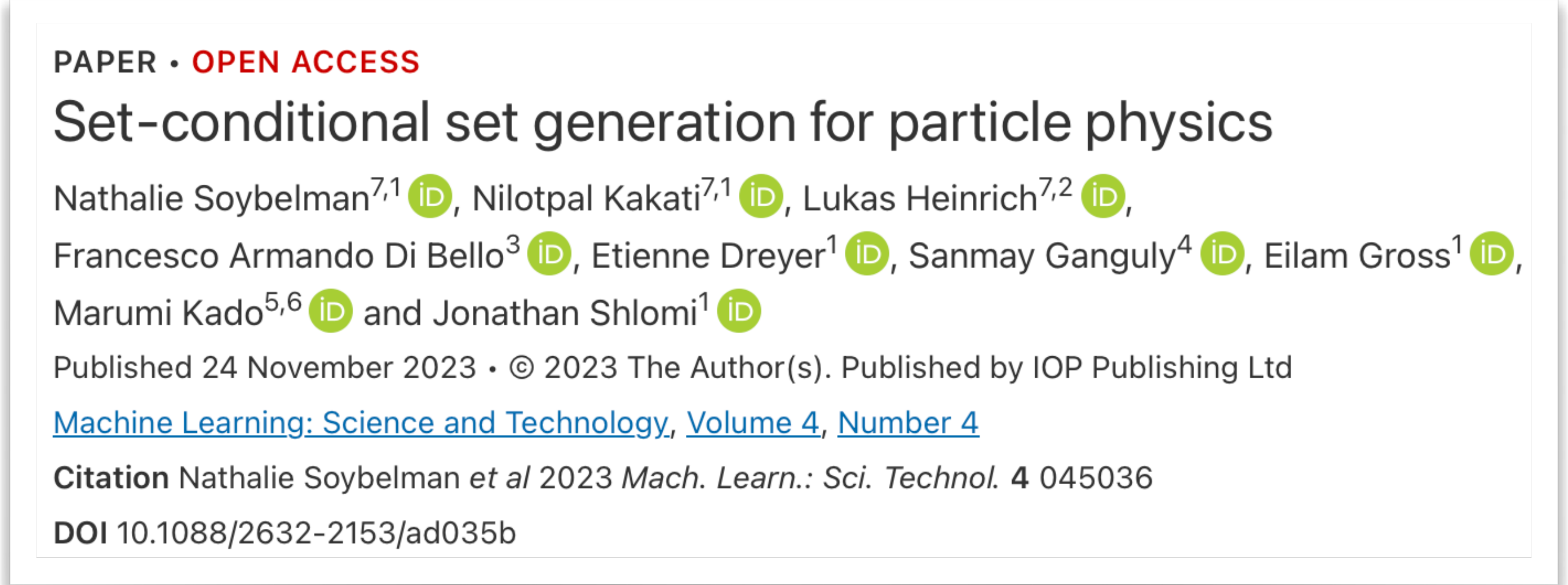
Jonathan Shlomi[1], Peter Battaglia[2] and Jean-Roch Vlimant[3]

Published 29 December 2020 • © 2020 The Author(s). Published by IOP Publishing Ltd

Machine Learning: Science and Technology, Volume 2, Number 2

Citation Jonathan Shlomi *et al* 2021 *Mach. Learn.: Sci. Technol.* **2** 021001

DOI 10.1088/2632-2153/abbf9a



PAPER • OPEN ACCESS

Set-conditional set generation for particle physics

Nathalie Soybelman[7,1], Nilotpal Kakati[7,1], Lukas Heinrich[7,2], Francesco Armando Di Bello[3], Etienne Dreyer[1], Sanmay Ganguly[4], Eilam Gross[1], Marumi Kado[5,6] and Jonathan Shlomi[1]

Published 24 November 2023 • © 2023 The Author(s). Published by IOP Publishing Ltd

Machine Learning: Science and Technology, Volume 4, Number 4

Citation Nathalie Soybelman *et al* 2023 *Mach. Learn.: Sci. Technol.* **4** 045036

DOI 10.1088/2632-2153/ad035b

MACHINE
LEARNING
Science and Technology

Web: https://iopscience.org/mlst   Follow: @MLSTjournal

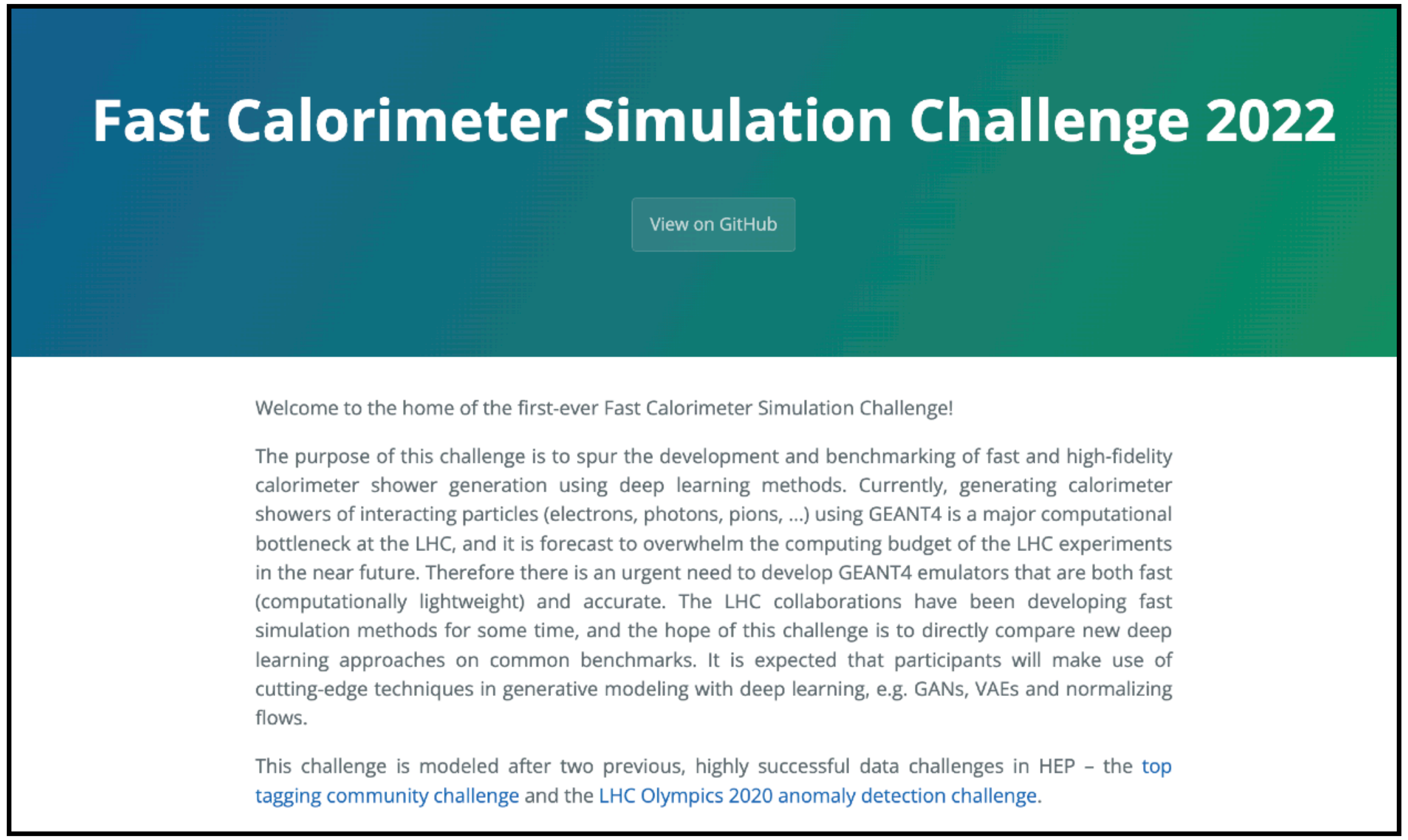
Challenges, Benchmarks, and Open Data

Higgs challenge

the HiggsML challenge

May to September 2014

When High Energy Physics meets Machine Learning

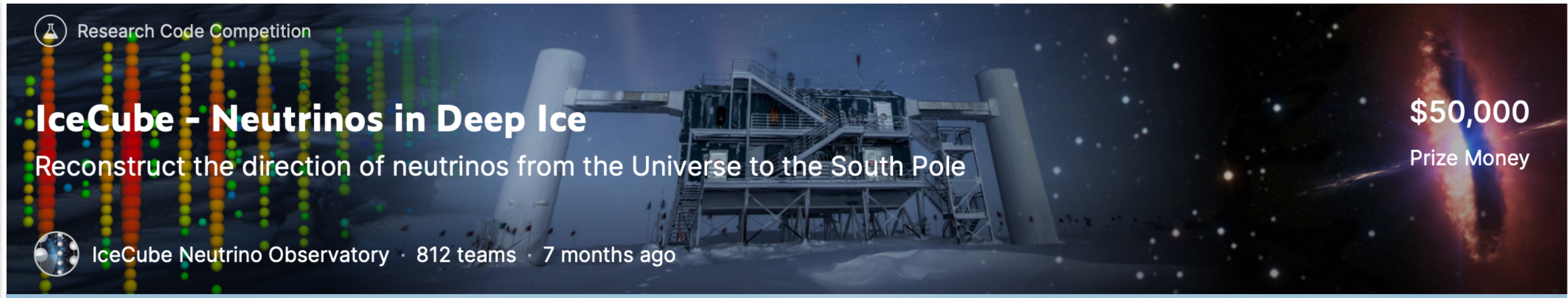


Fast Calorimeter Simulation Challenge 2022

Welcome to the home of the first-ever Fast Calorimeter Simulation Challenge!

The purpose of this challenge is to spur the development and benchmarking of fast and high-fidelity calorimeter shower generation using deep learning methods. Currently, generating calorimeter showers of interacting particles (electrons, photons, pions, ...) using GEANT4 is a major computational bottleneck at the LHC, and it is forecast to overwhelm the computing budget of the LHC experiments in the near future. Therefore there is an urgent need to develop GEANT4 emulators that are both fast (computationally lightweight) and accurate. The LHC collaborations have been developing fast simulation methods for some time, and the hope of this challenge is to directly compare new deep learning approaches on common benchmarks. It is expected that participants will make use of cutting-edge techniques in generative modeling with deep learning, e.g. GANs, VAEs and normalizing flows.

This challenge is modeled after two previous, highly successful data challenges in HEP – the top tagging community challenge and the LHC Olympics 2020 anomaly detection challenge.



Research Code Competition

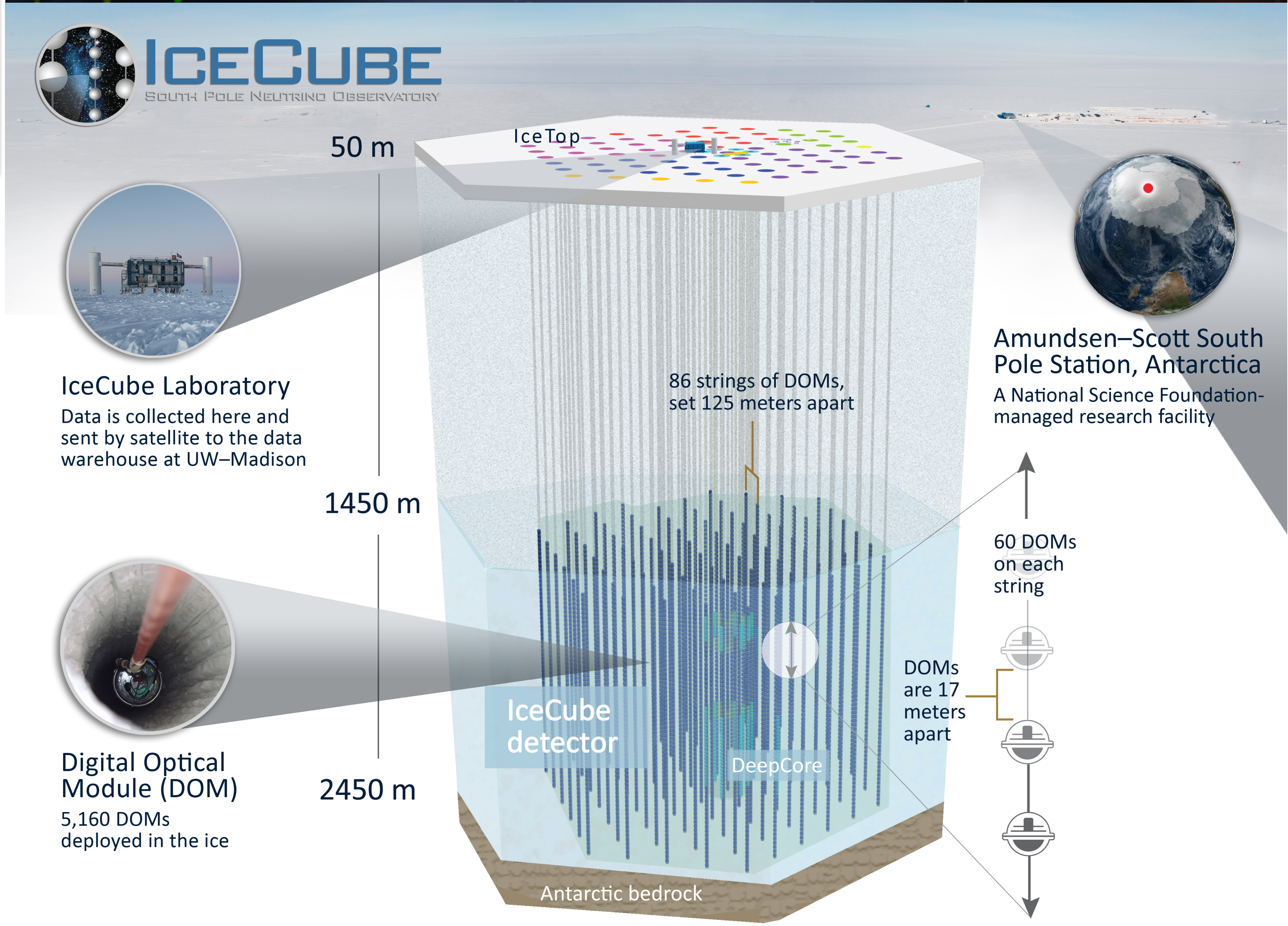
IceCube - Neutrinos in Deep Ice

Reconstruct the direction of neutrinos from the Universe to the South Pole

IceCube Neutrino Observatory · 812 teams · 7 months ago

$50,000

Prize Money



Collaborators + Many More


Gilles Louppe
U. Liège


Kyunghyun Cho


Joan Bruna


Sebastian Macaluso


Matthew Drnevich


Lauren Greenspan


Miles Cranmer


@HServiansky


Nimrod Segol


Johann Brehmer


Tilman Plehn


Lukas Heinrich


Peter Battaglia


Alvaro Sanchez Gonzalez


Shirley Ho

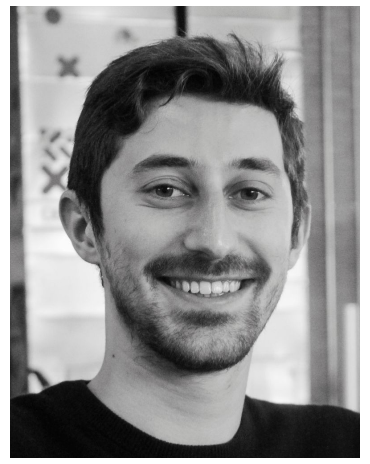

Irina Espejo


@HaggaiMaron


Yaron Lipman


Peter Sadowski

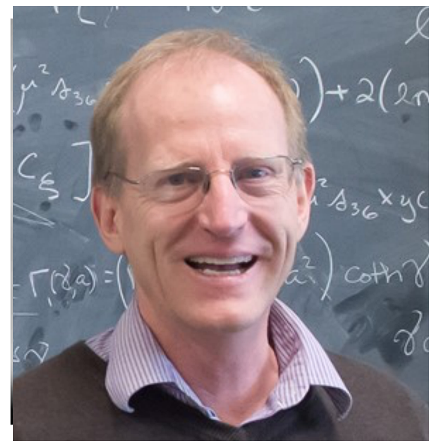

Daniel Whiteson


Michael Albergo

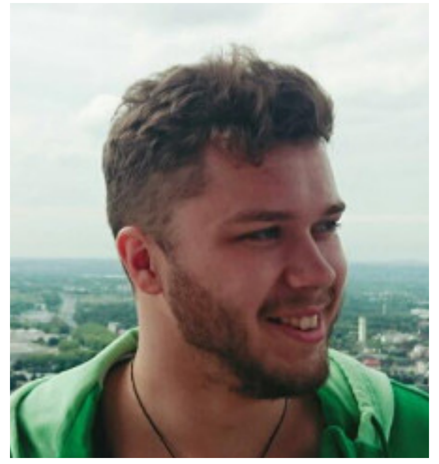

Danilo Rezende


Atılım Güneş Baydin


Garrett Merz


Lance Dixon


Tianji Cai


Niklas Nolte


Francois Charton


Phiala Shanahan


William Detmold


Gurtej Kanwar


Dan Hackett


Denis Boyda


Sébastien Racanière


Siddharth Mishra-Sharma


Patrick Rieck


Matthias Wilhelm

Fundamental Physics has a compelling story to tell and an important role in AI4Science.

… and it is critical that we tell it!

EuCAIF can help

Thank you!

Questions?