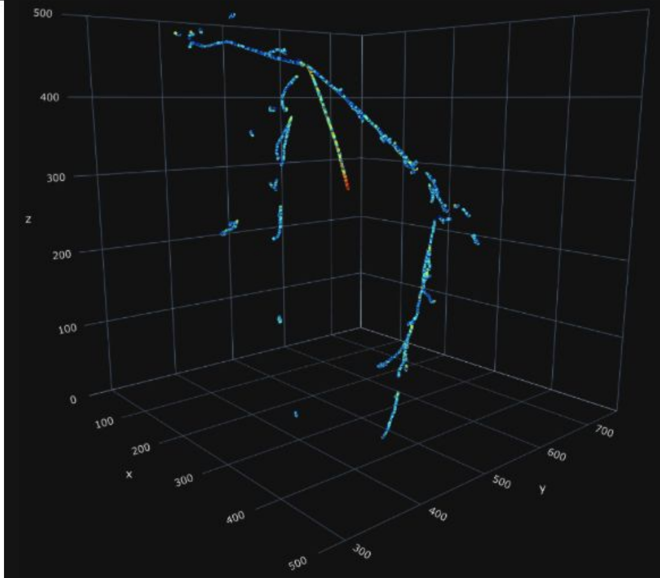
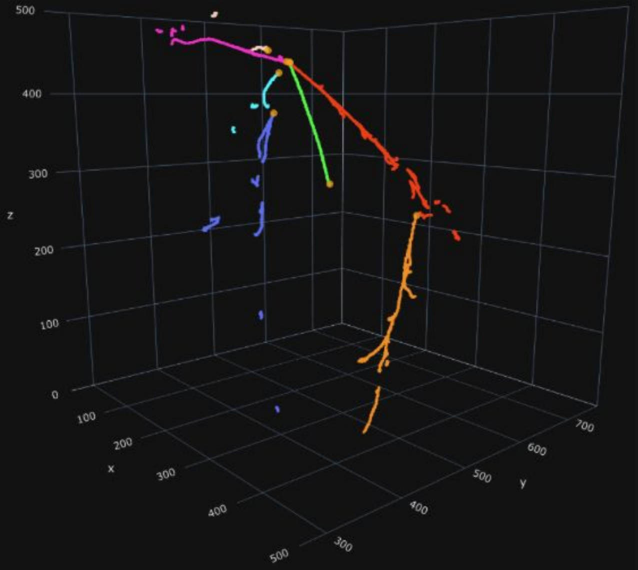
Machine Learning
Science Discoveries

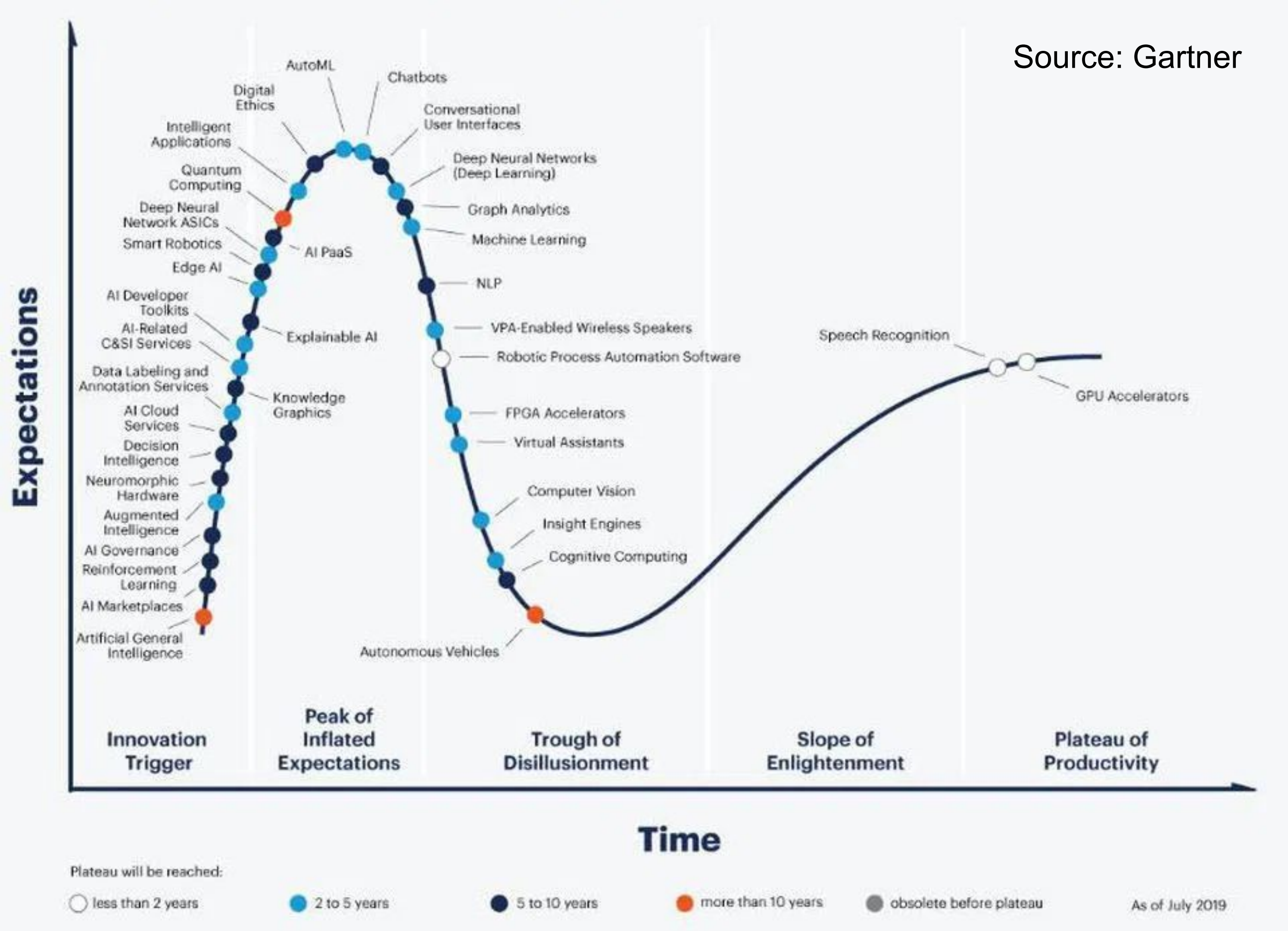




Original image credit: xkcd

Kazuhiro Terao (SLAC)
TRIUMF Science Week 2020

“... we would ask you to speak about the future of machine learning and its potential applications in the field over the coming 20 years.” - organizers

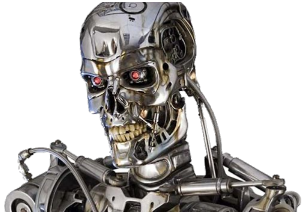


Many relevant for us!

- General AI
- Neuromorphic Hardware
- Explainable AI
- Edge AI
- Quantum Computing
- Conversational UI
- DNN
- Graph Analytics
- NLP
- FPGA accelerators
- Computer Vision
- GPU accelerators

… + more :)

SCIENCE!



"... we would ask you to speak about the future of machine learning and its potential applications in the field over the coming 20 years." - organizers

Goal: make the *hype cycle* for ML in Science

How: look at science and technology frontiers for ML application development in science

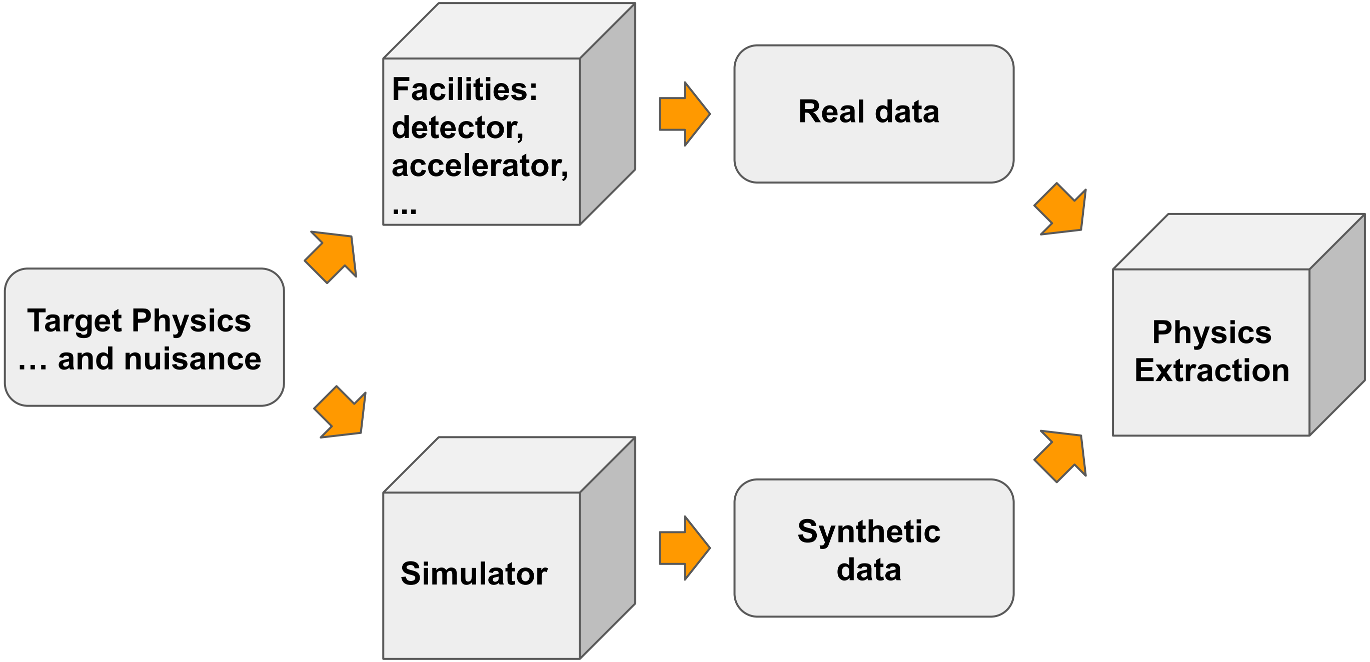
Impactful Scientific Applications in 5, 10, and 20 years

SLAC

Outline

1. Machine learning for scientific experiments
2. Advancement in eco-systems around ML
3. Hype cycle in science

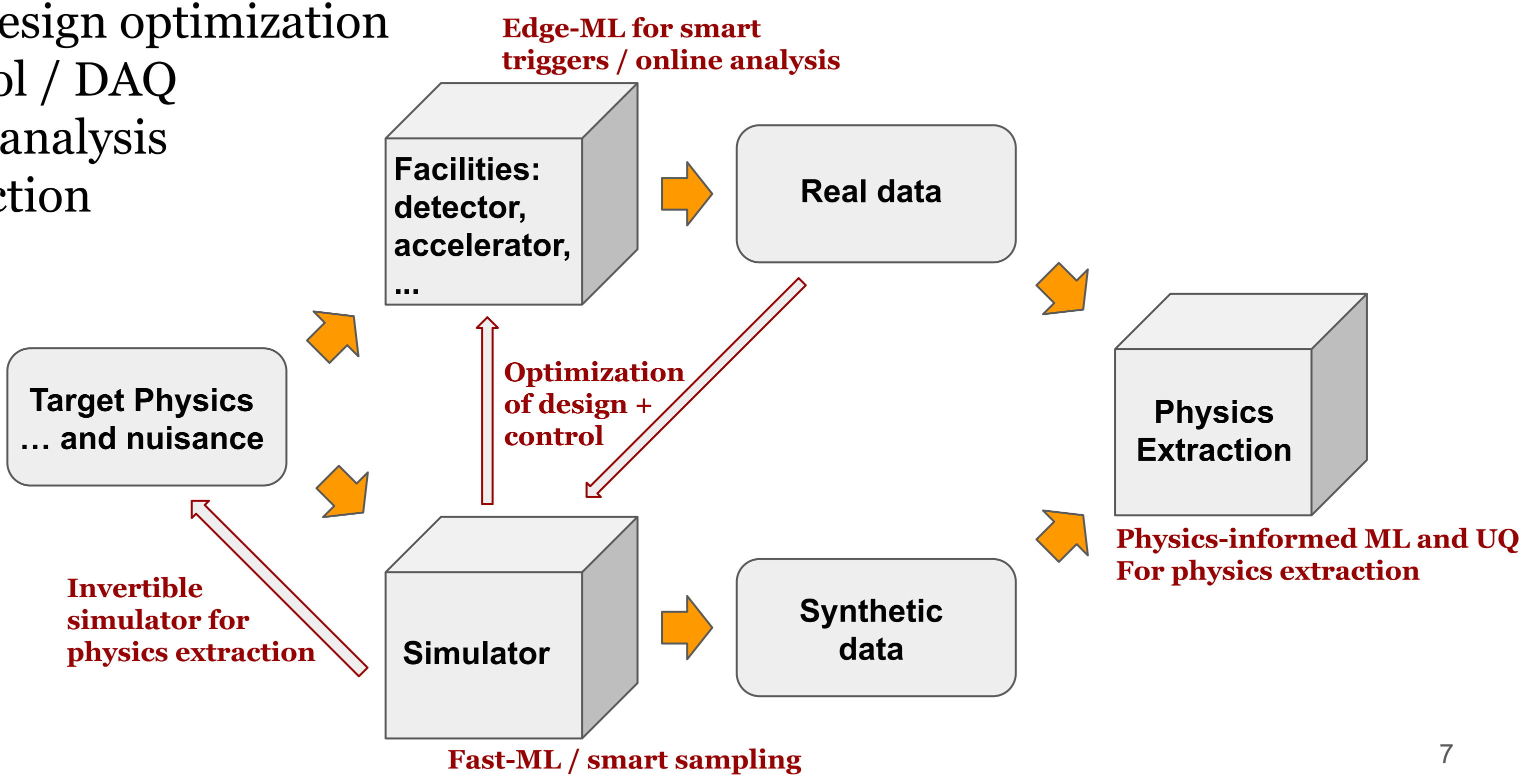
Landscape: where ML applied?



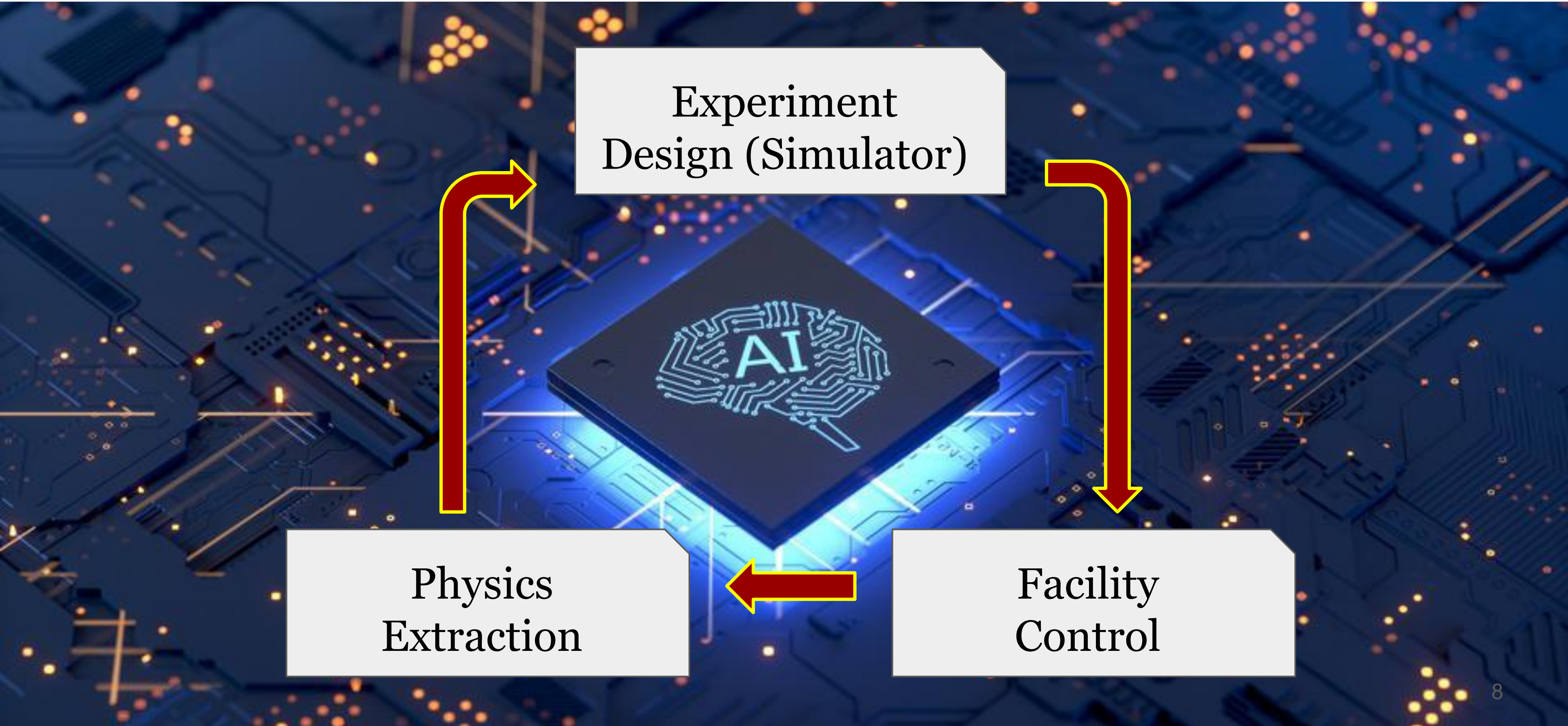
Machine Learning Discoveries
ML for Science Experiments

Landscape: where ML applied?

- Experiment design optimization
- Facility control / DAQ
- Simulation & analysis
- Physics extraction



Machine Learning Discoveries
ML for Science Experiments



Machine Learning Discoveries

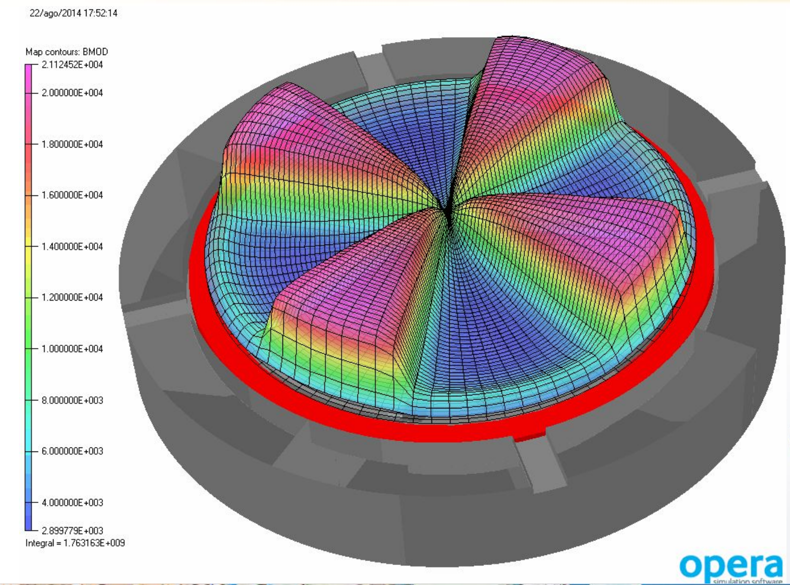
Experiment Design Optimization

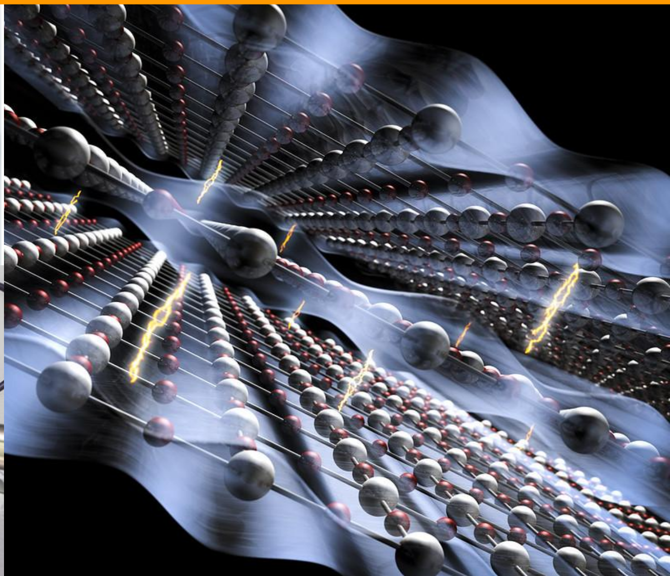
SLAC

Ingredient: physics modeling = simulator

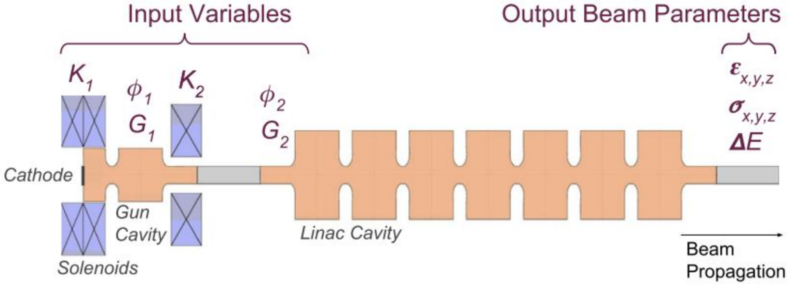
Goal: optimize the configuration parameters to optimize an objective function for design metrics

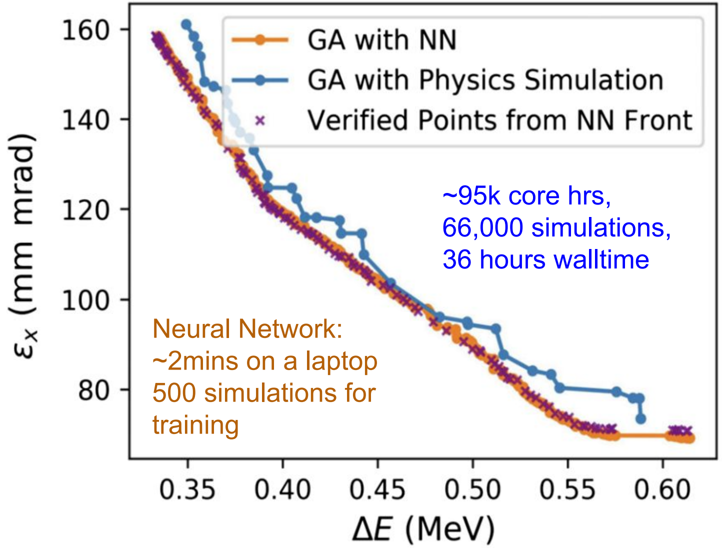
Challenge: simulator complex and expensive





Machine Learning Discoveries
Experiment Design Optimization





- NN surrogate: fast cheap simulation
  - Genetic algorithm (GA) with NN surrogate v.s. physics simulator for accelerator design optimization study.
  - 36 hours on HPC with a physics simulator
  - 2 minutes on laptop with NN surrogate
    - +17 minutes simulation + training
  - Phys. Rev. Accel. Beams 23, 044601 (2020)

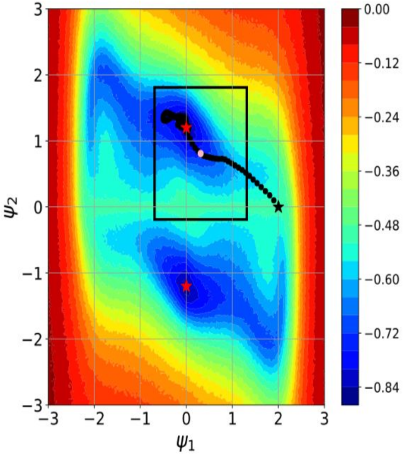
Machine learning for orders of magnitude speedup in multiobjective optimization of particle accelerator systems

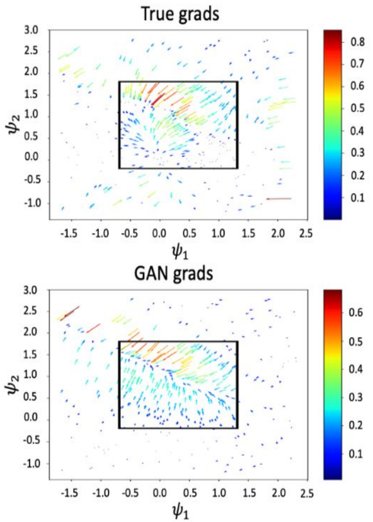
Auralee Edelen,[1,*] Nicole Neveu,[1] Matthias Frey,[2] Yannick Huber,[2] Christopher Mayes,[1] and Andreas Adelmann[2,†]

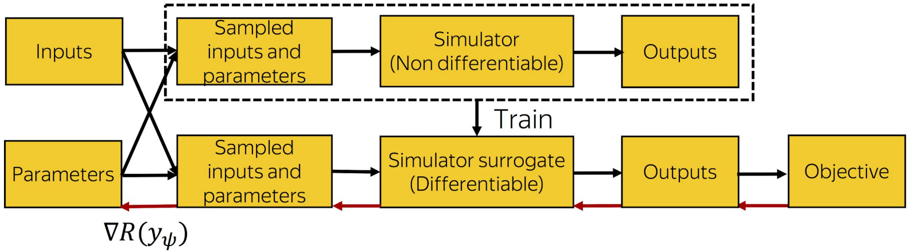
[1]*SLAC National Laboratory, Menlo Park, 94025 California, USA*
[2]*Paul Scherrer Institut, 5232 Villigen, Switzerland*

(Received 2 October 2019; accepted 24 January 2020; published 8 April 2020)

Machine Learning Discoveries
Experiment Design Optimization







- NN surrogate: gradient optimization
  - Generative NN surrogate to approximate the stochastic gradient of true simulator (non-differentiable), which enables direct optimization using back-propagation
  - arXiv:2002.04632... promising initial work!
    - ... with a comparison to other methods (Bayesian optimization using Gaussian Process, etc.)
  - Future directions: how does it scale for a large system? Would it be stable?
    - Could we make differentiable simulator?

Machine Learning Discoveries
Experiment Design Optimization

Ingredient: physics modeling = simulator

Goal: optimize the configuration parameters to optimize an objective function for design metrics

Challenge: simulator complex and expensive

Take aways

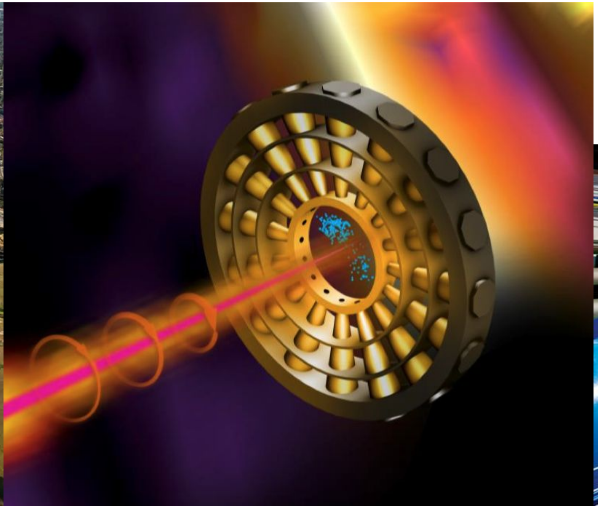
- Stochastic simulation (e.g. particle scattering) = non-differentiable likelihood often intractable to use directly for optimization
- ML surrogates for black-box (simulator) optimization as an alternative
  - Also applicable: Bayesian optimization using GP (later), likelihood free inference (later), etc. but less used for design optimization
- Let’s take a good design = less $$ more science!

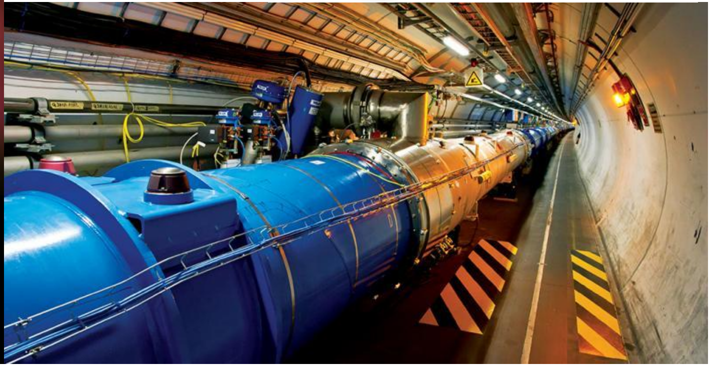
Machine Learning Discoveries
Facility operation and data taking

Ingredients: accelerator, detector, DAQ, monitoring systems

Goal: improve the detector/facility operations and data quality

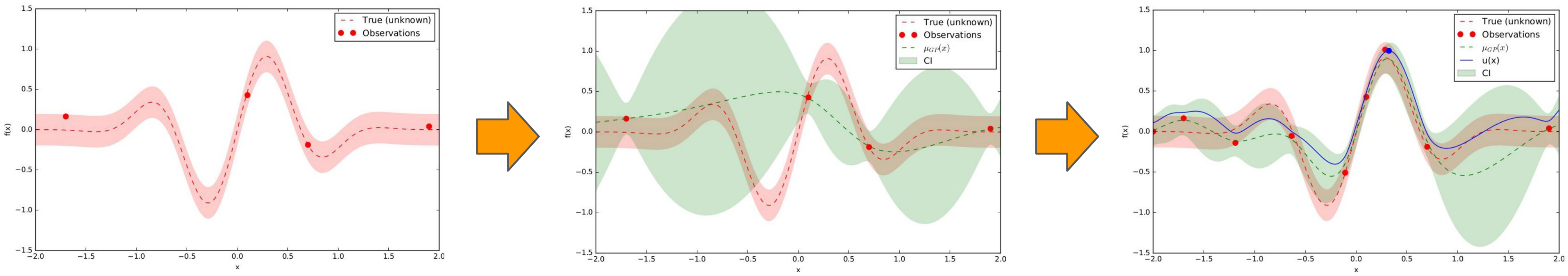
Challenge: active systems = speed and efficiency are the keys!





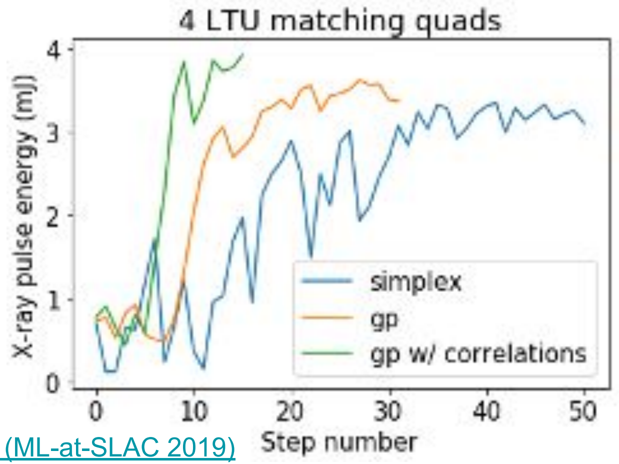
Machine Learning Discoveries
Accelerator operations

Bayesian optimization (e.g. Gaussian Process) for efficient tuning



Figures courtesy of [Gilles Louppe (PyData 2017)]

- Tuning of quadrupole magnet at LCLS
  - Probabilistic model = interpretability
  - GP v.s. “Hand-tuning” = x 2~3 times faster
    - [Phys. Rev. Lett. 124, 124801]
  - LCLS “hand-tuning” (not only quadrupole) time ~400 hours/year, $12M!



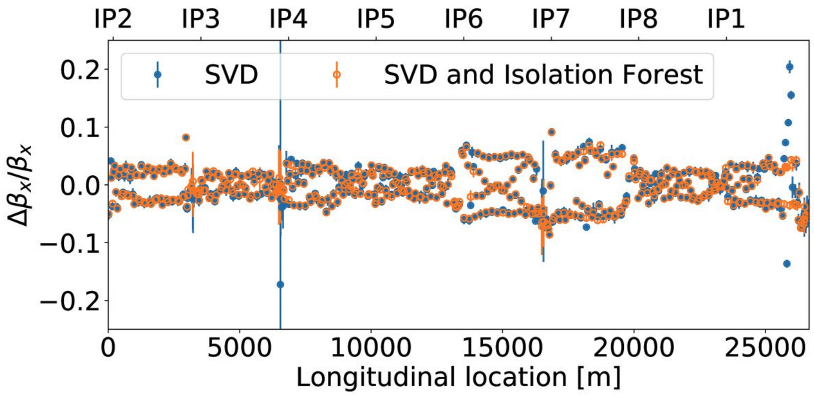
Courtesy of [Joe Duris (ML-at-SLAC 2019)]

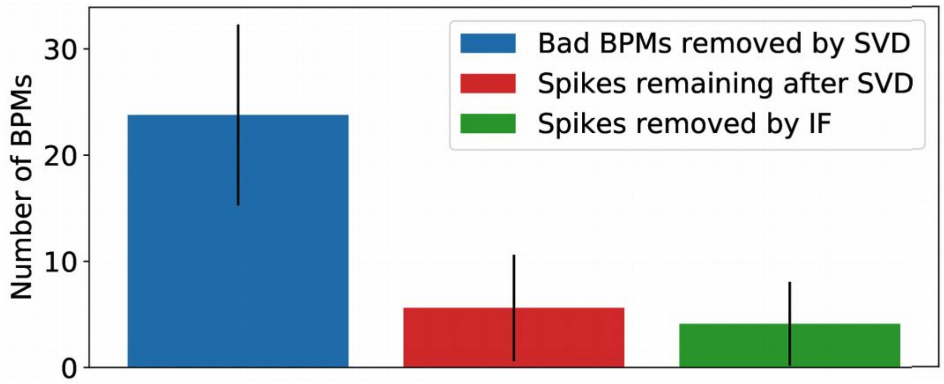
Machine Learning Discoveries
Accelerator operations

SLAC

Anomaly detection for finding false beam position monitor (BPM) signals

- Standard method (Single Value Decomposition = SVD) removes most of faulty BPM measurements at LHC but not all!
- Isolation Forest (IF, unsupervised method using binary trees) removes the majority
- [Elena Fol et al. ICPF 2019]()





Machine Learning Discoveries
Online triggers + GPUs

High data throughput **GPU-based trigger** system (LHCb)

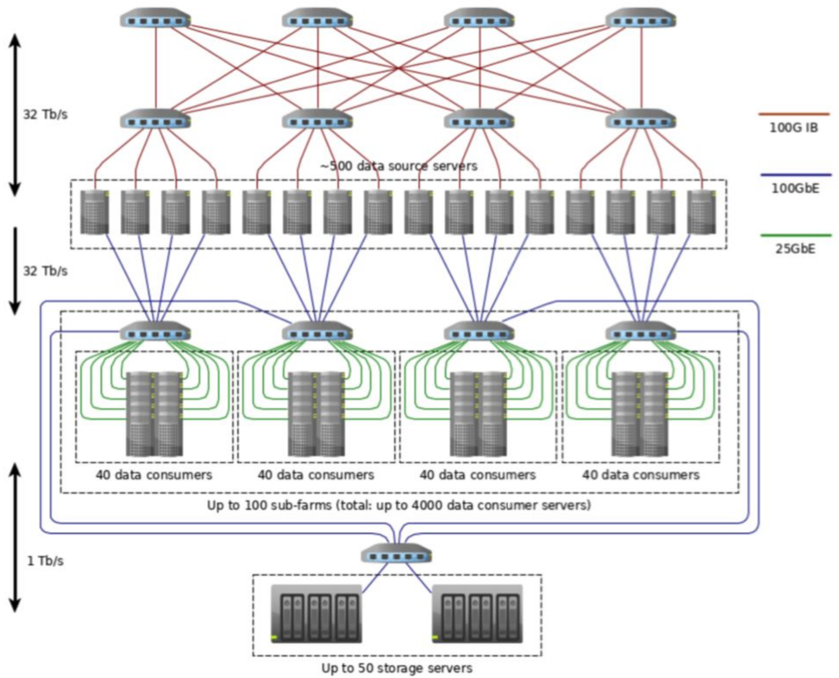


Image from [LHCb TDR]

Computing and Software for Big Science
This is a post-peer-review, pre-copyedit version of this article.
The final authenticated version is available online at: https://doi.org/10.1007/s41781-020-00039-7

Allen: A high level trigger on GPUs for LHCb

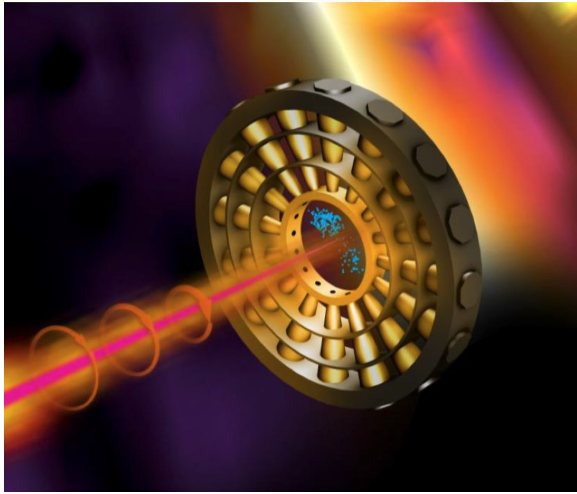
R. Aaij* · J. Albrecht · M. Belous · P. Billoir · T. Boettcher · A. Brea Rodríguez · D. vom Bruch* · D. H. Cámpora Pérez* · A. Casais Vidal · D. C. Craik · P. Fernandez Declara · L. Funke · V. V. Gligorov · B. Jashal · N. Kazeev · D. Martínez Santos · F. Pisani · D. Pliushchenko · S. Popov · R. Quagliani · M. Rangel · F. Reiss · C. Sánchez Mayordomo · R. Schwemmer · M. Sokoloff · H. Stevens · A. Ustyuzhanin · X. Vilasís Cardona · M. Williams

- [arXiv:1912.09161]
- 500 GPUs for collision rate @ 30 MHz = ~40 Tb/s
- Key element: data bandwidth
  - FPGA (next slide) for predictable latency

Machine Learning Discoveries
Online triggers + FPGAs

SLAC

ML on FPGA @ Linear Coherent Light Source



- Data rate 20 - 1200 GB/s at 1 MHz beam rate
  - 10 kHz at early LCLS-II
- Pipelined MLP on FPGA = 19.3 micro-seconds latency @ 77 kHz throughput, more architectures tested (see a talk by Audrey T. and Ryan C. at DANCE-ML 2020)

hls4ml

HLS4ML = **(Physicists + ML)/FPGA**

- Automatic translation of open-source ML model to HLS + compile on FPGA
- Meant to be generic, reusable framework

Machine Learning Discoveries
Facility operation and data taking

Ingredients: accelerator, detector, DAQ, monitoring systems

Goal: improve the detector/facility operations and data quality

Challenge: active systems = speed and efficiency are the keys!

Take aways

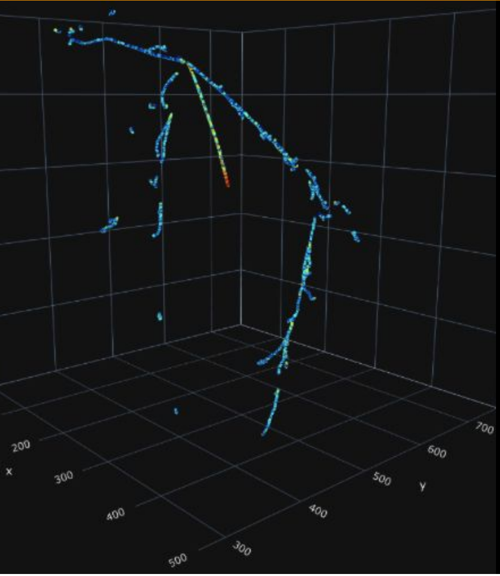
- Probabilistic models for operations support
- Efficient sampling (Bayesian optimization) for fast turn-around
- Anomaly detection
- Edge/Fast-ML to bring high level analysis to the detector
- Active learning: could we learn from data online?

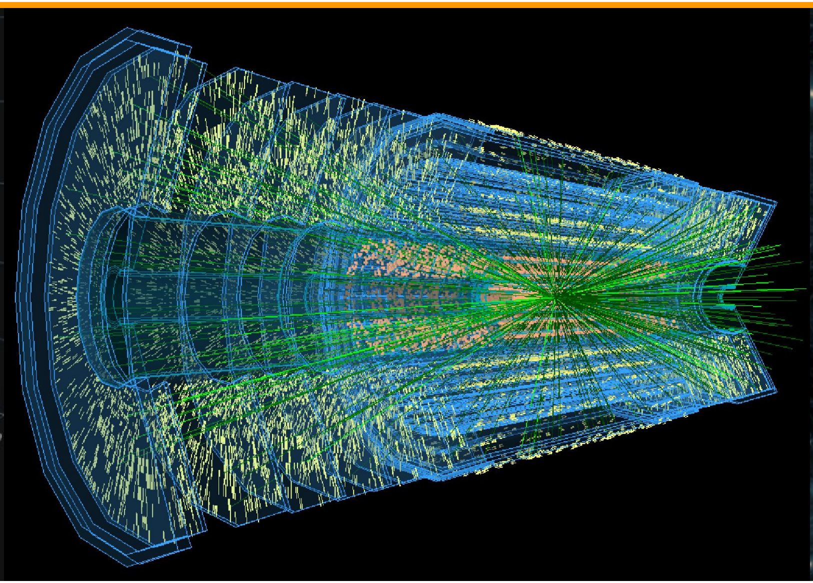
Machine Learning Discoveries
Data analysis & physics inference

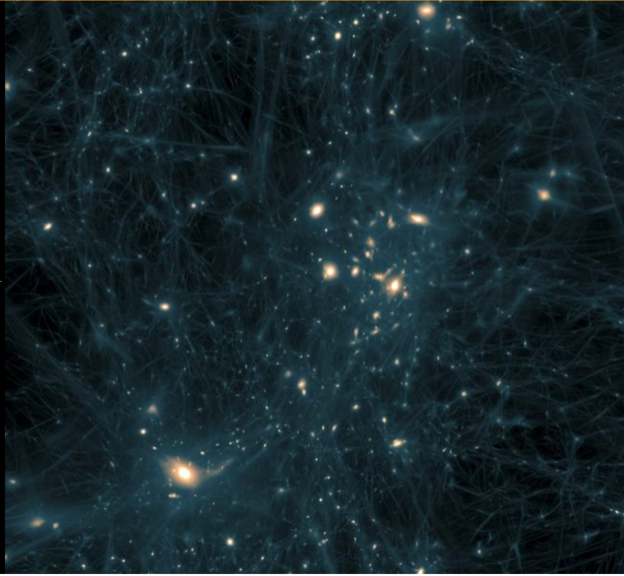
Ingredients: large, multi-modal detector big data

Goal: extract physics signal

Challenge: irregular data structure, interpretable high quality analysis





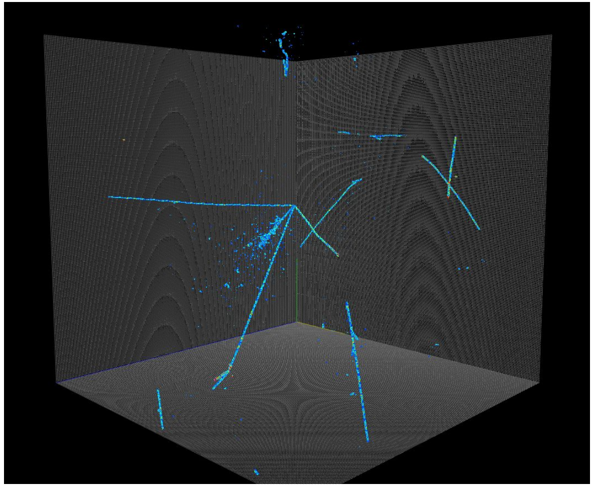


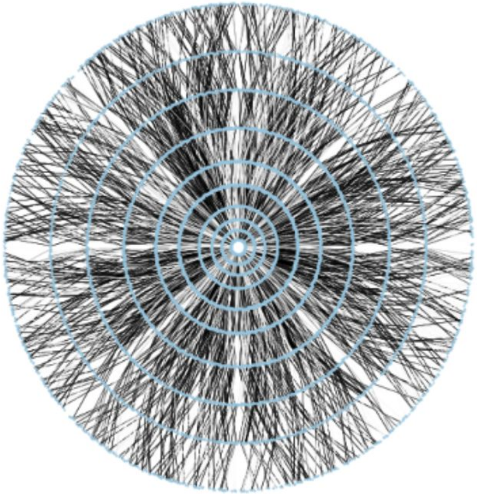
Machine Learning Discoveries
Science-domain ML

- **Data structure**: sparse images, point cloud data, detector geometry

Scalable Deep Convolutional Neural Networks for Sparse, Locally Dense Liquid Argon Time Projection Chamber Data

Laura Dominé[1,2] and Kazuhiro Terao[2]
(on behalf of the DeepLearnPhysics Collaboration)*
[1] *Stanford University, Stanford, CA, 94305, USA*
[2] *SLAC National Accelerator Laboratory, Menlo Park, CA, 94025, USA*





Graph Neural Networks for Particle Reconstruction in High Energy Physics detectors

Xiangyang Ju, Steven Farrell, Paolo Calafiura, Daniel Murnane, Prabhat
Lawrence Berkeley National Laboratory
Berkeley, CA
xju@lbl.gov

Lindsey Gray, Thomas Klijnsma, Kevin Pedro, Giuseppe Cerati, Jim Kowalkowski, Gabriel Perdue, Panagiotis Spentzouris, Nhan Tran
Fermi National Accelerator Laboratory
Batavia, IL

Jean-Roch Vlimant, Alexander Zlokapa, Joosep Pata, Maria Spiropulu
California Institute of Technology
Pasadena, CA

Sitong An
CERN, Geneva, Switzerland &
Carnegie Mellon University, Pittsburgh, PA

Adam Aurisano, Jeremy Hewes
University of Cincinnati
Cincinnati, OH

Aristeidis Tsaris
Oak Ridge National Laboratory
Oak Ridge, TN

Kazuhiro Terao, Tracy Usher
SLAC National Accelerator Laboratory
Menlo Park, CA

Machine Learning Discoveries
Science-domain ML

- **Symmetry**: cylindrical/spherical detector, ML with SU(3), etc.

Published as a conference paper at ICLR 2018

HexaConv

Emiel Hoogeboom*, **Jorn W.T. Peters*** **& Taco S. Cohen**
University of Amsterdam
{e.hoogeboom,j.w.t.peters,t.s.cohen}@uva.nl

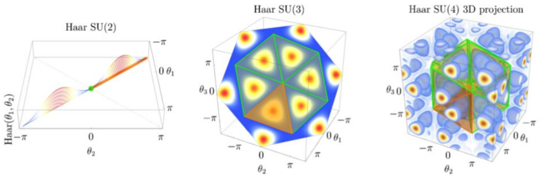
Max Welling
University of Amsterdam & CIFAR
m.welling@uva.nl

Lorentz Group Equivariant Neural Network for Particle Physics

Alexander Bogatskiy [1] **Brandon Anderson** [2,3] **Jan T. Offermann** [1] **Marwah Roussi** [1] **David W. Miller** [1,4]
Risi Kondor [2,5,6]

Sampling using $\mathrm{SU}(N)$ gauge equivariant flows

Denis Boyda,[1,*] Gurtej Kanwar,[1,†] Sébastien Racanière,[2,‡] Danilo Jimenez Rezende,[2,§]
Michael S. Albergo,[3] Kyle Cranmer,[3] Daniel C. Hackett,[1] and Phiala E. Shanahan[1]
[1] *Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*
[2] *DeepMind, London, UK*
[3] *Center for Cosmology and Particle Physics, New York University, New York, NY 10003, USA*
(Dated: August 13, 2020)



Published as a conference paper at ICLR 2018

Spherical CNNs

Taco S. Cohen* University of Amsterdam
Mario Geiger* EPFL
Jonas Köhler* University of Amsterdam

Max Welling
University of Amsterdam & CIFAR

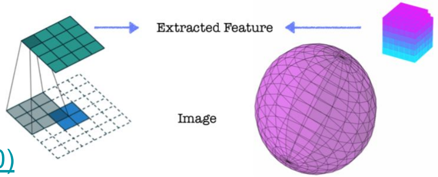
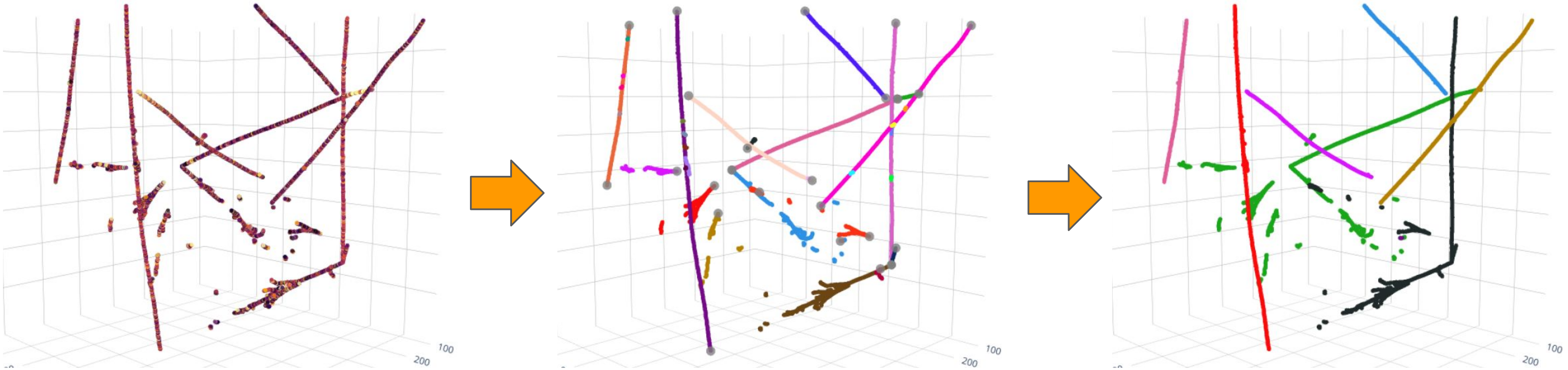


Image from
[Aobo Li's talk (NPML 2020)]

Machine Learning Discoveries
Science-domain ML

- **Interpretability**: hierarchical, compositional structure



Point Proposal Network for Reconstructing 3D Particle Endpoints with Sub-Pixel Precision in Liquid Argon Time Projection Chambers

Laura Dominé,[3,*] Pierre Côte de Soux,[2] François Drielsma,[1] Dae Heun Koh,[3]
Ran Itay,[1] Qing Lin,[1] Kazuhiro Terao,[1] Ka Vang Tsang,[1] and Tracy L. Usher[1]
(on behalf of the DeepLearnPhysics Collaboration)

Scalable, Proposal-free Instance Segmentation Network for 3D Pixel Clustering and Particle Trajectory Reconstruction in Liquid Argon Time Projection Chambers

Dae Heun Koh,[3,*] Pierre Côte de Soux,[2] Laura Dominé,[3] François Drielsma,[1]
Ran Itay,[1] Qing Lin,[1] Kazuhiro Terao,[1] Ka Vang Tsang,[1] and Tracy L. Usher[1]
(on behalf of the DeepLearnPhysics Collaboration)

ML for end-to-end multi-stage reconstruction enforce inductive bias and make analysis output interpretable with hierarchical/sequential evidence finding

Clustering of Electromagnetic Showers and Particle Interactions with Graph Neural Networks in Liquid Argon Time Projection Chambers Data

François Drielsma,[1,*] Qing Lin,[1] Pierre Côte de Soux,[2] Laura Dominé,[3] Ran Itay,[1]
Dae Heun Koh,[3] Bradley J. Nelson,[2] Kazuhiro Terao,[1] Ka Vang Tsang,[1] and Tracy L. Usher[1]
(on behalf of the DeepLearnPhysics Collaboration)

Machine Learning Discoveries
Science-domain ML

- **Interpretability**: inductive bias / causal structure

ML4Jets

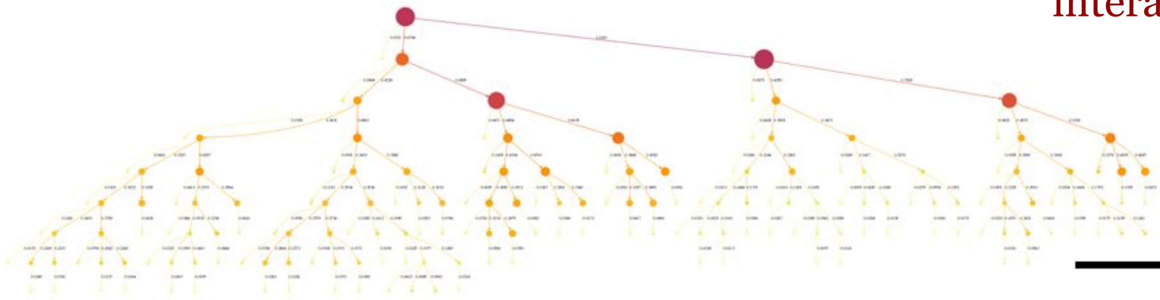


Image taken from IRIS-HEP

QCD-aware NNs incorporating interactions in trees and graphs

Neural Message Passing for Jet Physics

Isaac Henrion, Johann Brehmer, Joan Bruna, Kyunghun Cho, Kyle Cranmer
Center for Data Science
New York University
New York, NY 10012
{henrion*, johann.brehmer, bruna, kyunghyun, kyle.cranmer*}@nyu.edu

Gilles Louppe
Department of Computer Science
University of Liège
Belgium
g.louppe@ulg.ac.be

Gaspar Rochette
Department of Computer Science
École Normale Supérieure
Paris, France
gaspar.rochette@ens.fr

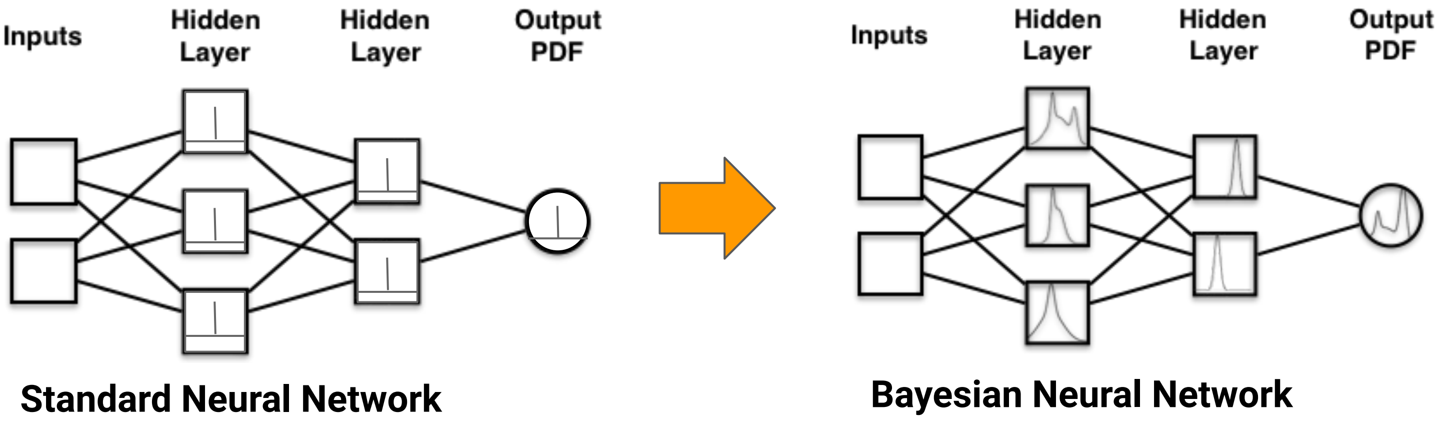
The Machine Learning Landscape of Top Taggers

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]

Machine Learning Discoveries
Science-domain ML

- **Interpretability**: uncertainty quantification = probabilistic approach
  - Model uncertainty and input systematic propagation
  - Systematic uncertainty for mismodeling of physics

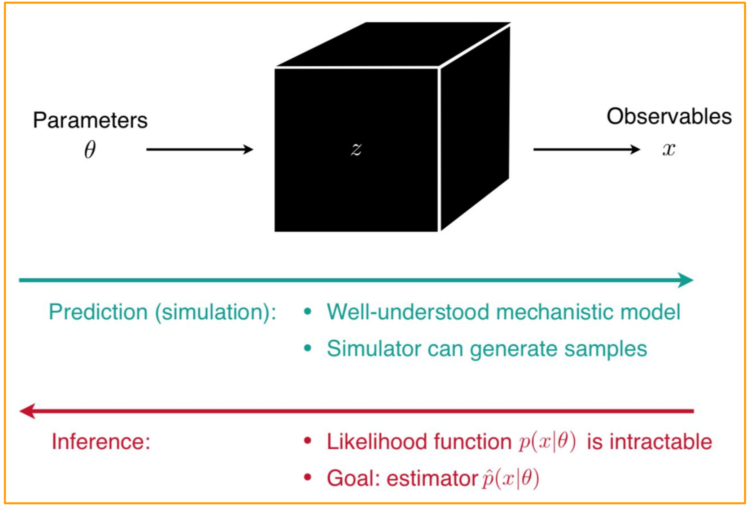
Natively designed methods: Bayesian NN, probabilistic programing, etc.



Machine Learning Discoveries
Science-domain ML

- **Interpretability**: uncertainty quantification = probabilistic approach
  - Model uncertainty and input systematic propagation
  - Systematic uncertainty for mismodeling of physics

Natively designed methods: Bayesian NN, probabilistic programing, etc.
… or solve inverse problem + simulator: Likelihood free inference



Monte Carlo Sampling

observed

what happened in simulation

$$p(x|\theta) = \boxed{\int dz}\, p(x, z|\theta)$$

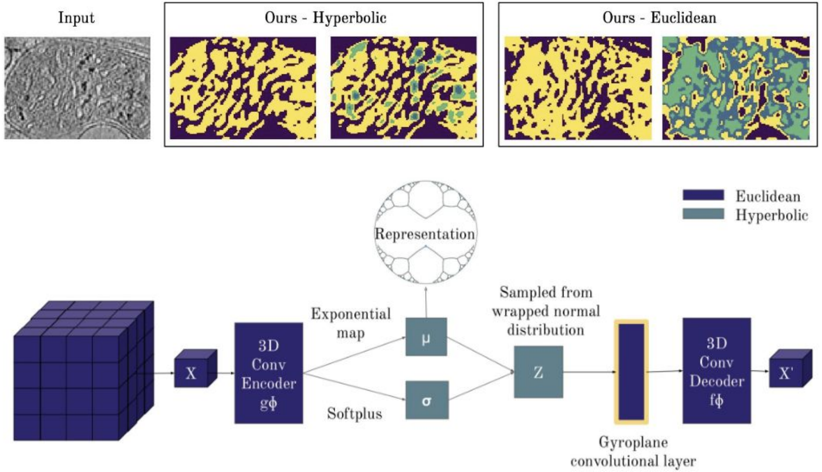
- Likelihood intractable: ways to approximate (e.g. generative model to “learn simulation”)

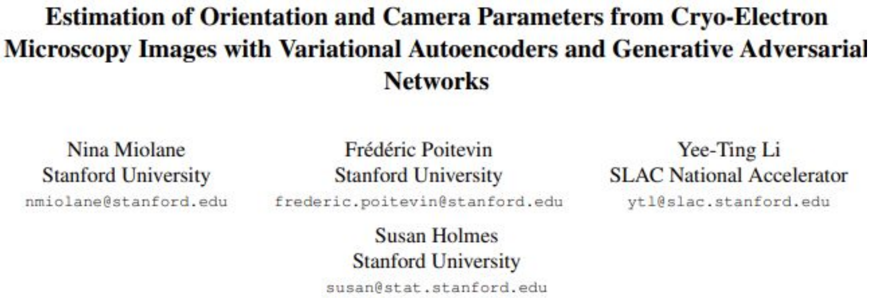
Slide by [Kyle Cranmer (Hammer & Nails 2019)]

Machine Learning Discoveries
Science-domain ML

- **Learning from data**: unsupervised generative models

"Learning Hyperbolic Representations for Unsupervised 3D Segmentation"
talk by [Joy Hsu (AI-at-SLAC seminar)]



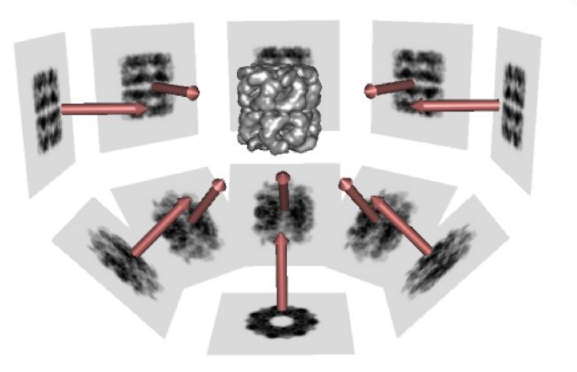

Estimation of Orientation and Camera Parameters from Cryo-Electron Microscopy Images with Variational Autoencoders and Generative Adversarial Networks

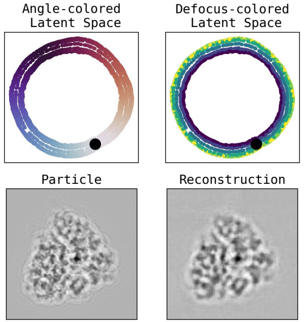
Nina Miolane
Stanford University
nmiolane@stanford.edu

Frédéric Poitevin
Stanford University
frederic.poitevin@stanford.edu

Yee-Ting Li
SLAC National Accelerator
ytl@slac.stanford.edu

Susan Holmes
Stanford University
susan@stat.stanford.edu





Images from a [talk by Frederic P. and Nina M. (DANCE-ML 2020)]

Data analysis & physics inference

Ingredients: large, multi-modal detector big data

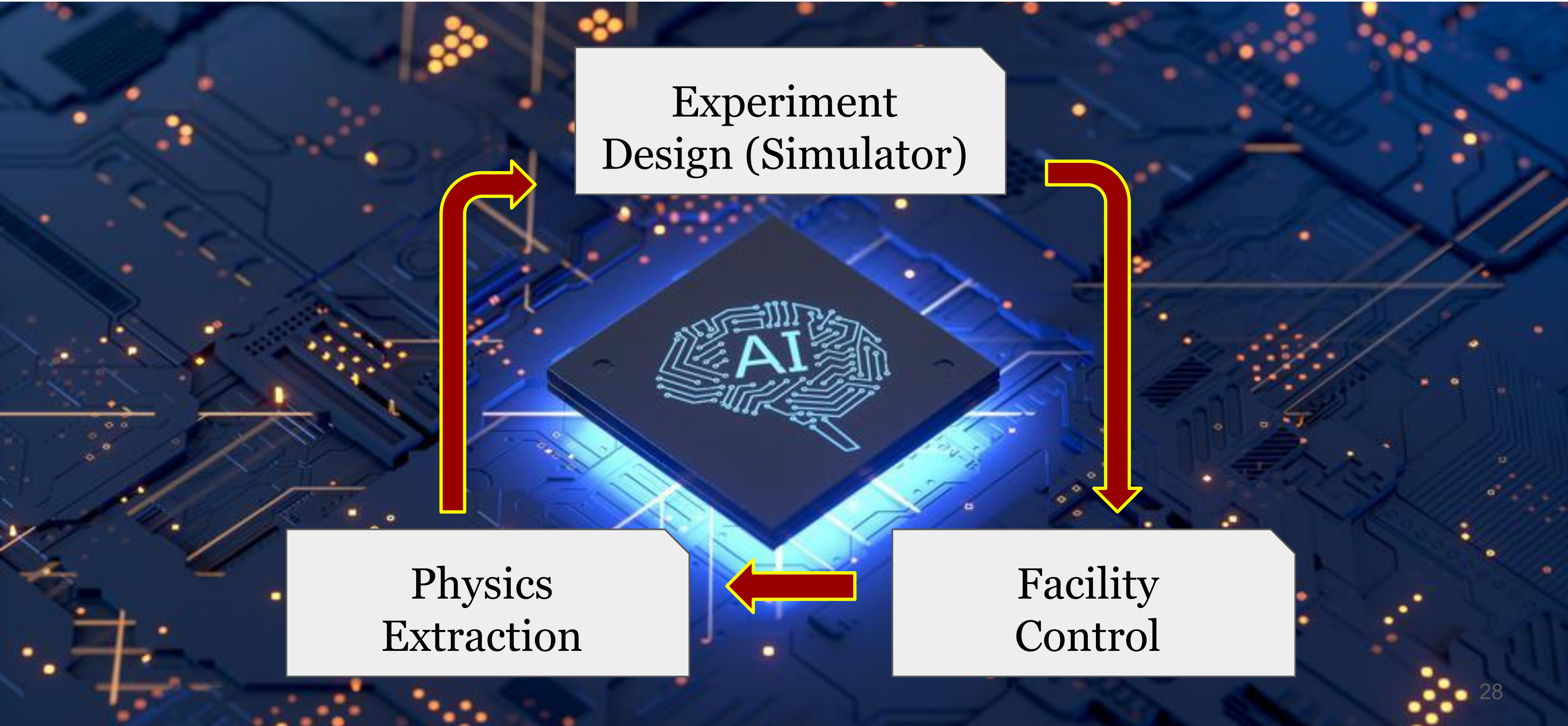
Goal: extract physics signal

Challenge: irregular data structure, interpretable high quality analysis

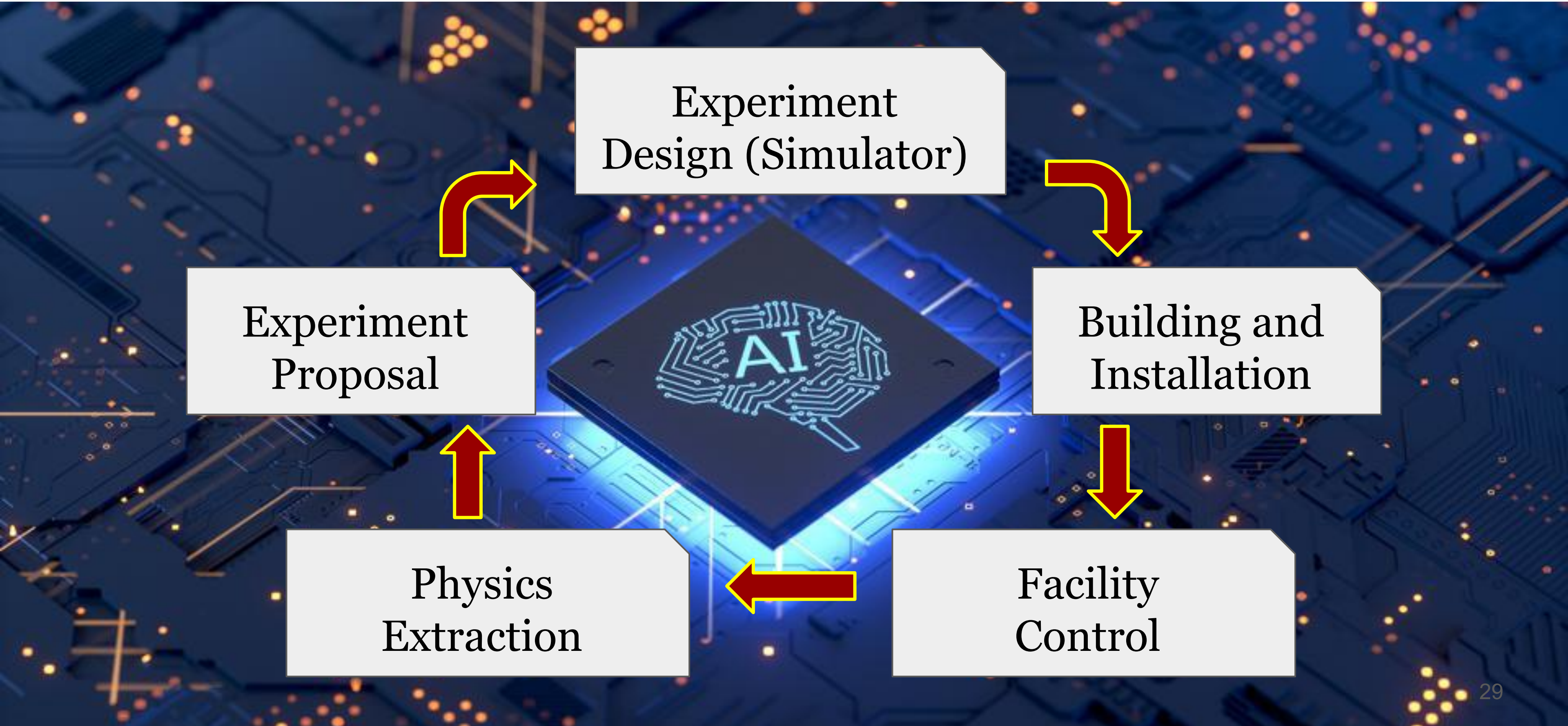
Take aways

- Science-informed ML is very active frontier of development
  - Domain-specific nature of data from multi-modal detectors
  - Enforcing symmetry and physics laws in architecture
- Interpretability
  - Enhance our knowledge: hierarchical, compositional, causal structure
  - Uncertainty estimation = intersection of ML and statistical methods

Machine Learning Discoveries
ML for Science Experiments



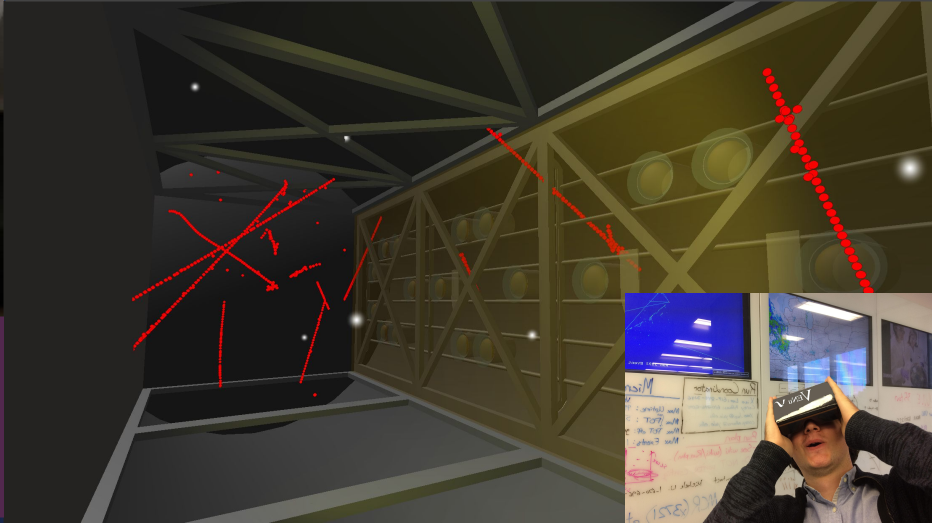
Machine Learning Discoveries
ML for Science Experiments



Machine Learning Discoveries
Eco-systems around Science

2019
Human-Centered
Artificial Intelligence
Symposium

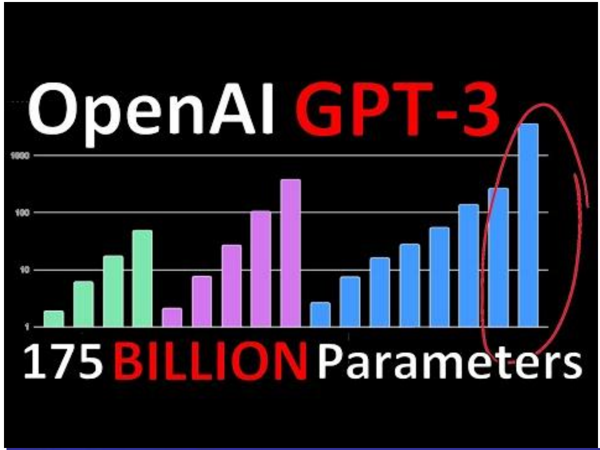
"Load all the wires from this event, loop over each one, find all the hits over the noise threshold, fit a gaussian to each, and save them as hits" and get a parallelized, compiler-friendly hitfinder out of the box.



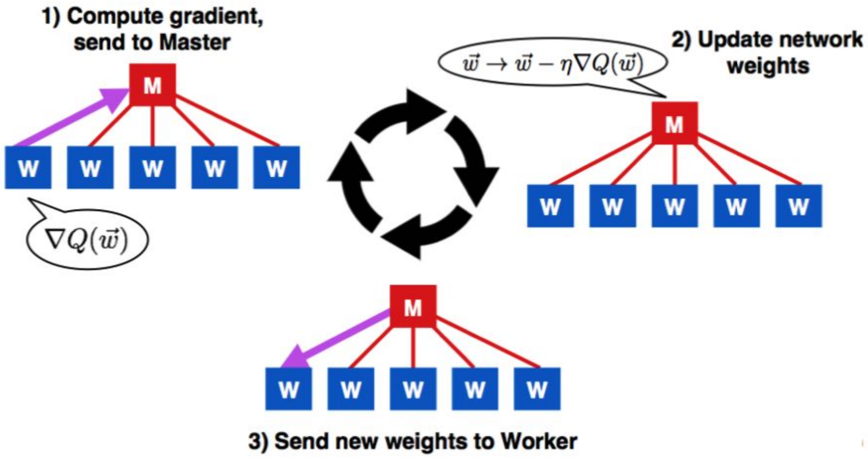
Machine Learning Discoveries
Eco-systems around ML

Computing

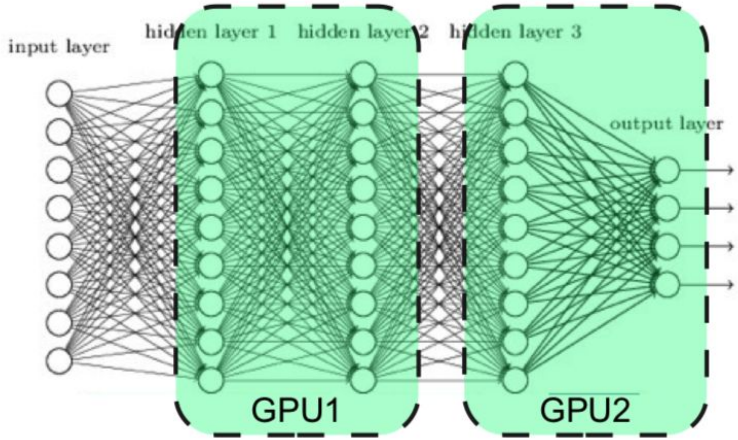
- Exa-scale HPC: next year!
  - GPU+CNN was only 8 years ago
  - ML on FPGA only a few years ago
  - … today's HPC on my laptop in 20 years?
  - "ASCI White" @ LLNL
    - 12.3 TFLOPS: fastest supercomputer (2002)
    - Today: NVIDIA 2080Ti 14 TFLOPS
- **Advancements solely by computing?**
  - Huge leap of performance without advancement in algorithm expected (e.g. OpenAI GPT-3)



Distributed Machine Learning



Data distributed training
arXiv:1712.05878



Model distributed training
T. Kurth et al. (SuperComputing 18)

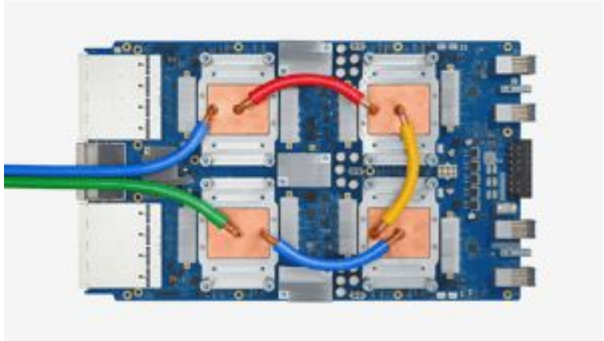
Machine Learning Discoveries
Eco-systems around ML

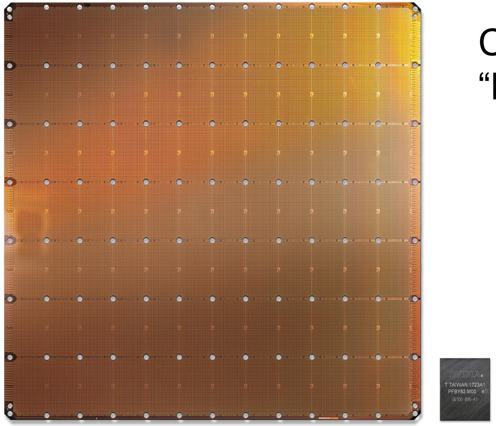
Evolving co-processors

- How do we design our “compute center”
- How to utilize LARGE #cores in a chip?
- How to benefit HUGE memory?

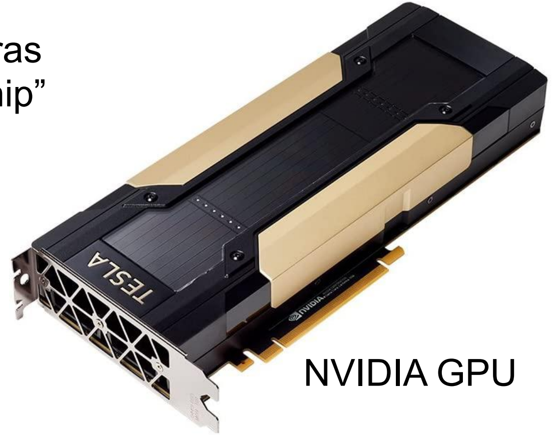
Xilink FPGA

Google TPU custom ASIC





Cerebras
“Big chip”



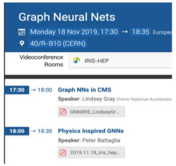
NVIDIA GPU

SLAC

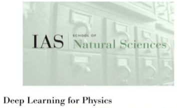
Eco-system

- How do we train new generations? No, how do we train ourselves?
    - Courses/workshops to be standardized in a wider community
- How to best foster academic/industrial research collaboration?
    - Funding support, open development with public benchmark data
- ML-in-Science = own field? (academic degrees, career path)

Microsoft sponsored

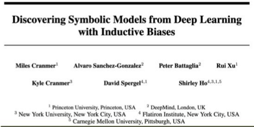


IRIS-HEP sponsored

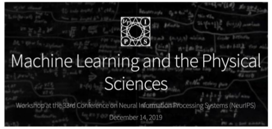


IAS/Princeton sponsored

Lectures Series
Various groups are organizing lectures where ML researchers present to physics groups (and vice versa)



Individual Collaborations
Many physicists are collaborating directly with industry researchers; some hold dual appointments across physics and CS



Events at ML Conferences
ML and Physical Sciences Workshop at NeurIPS (2017, 2019), resubmitted for 2020. Brings physics applications and physics for ML together.

MLHEP School
Weeklong training eveloped and sponsored by Yandex with collaboration from physicists (HEP focused)

The HEP Software Foundation facilitates cooperation and common efforts in High Energy Physics software and computing internationally.

HSF

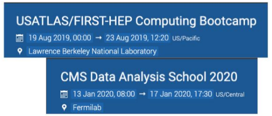
HEP Software Foundation Trainings
Materials for reproducible trainings, ML module in development (LHC focused)

Deep Learning for Science School

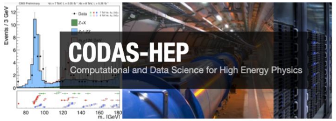
Webinar Series: July - September, 2020

Lawrence Berkeley National Laboratory, Berkeley, CA

Deep Learning for Science School
5 day training from ML experts and discussion on use in scientific research (General Science focused)



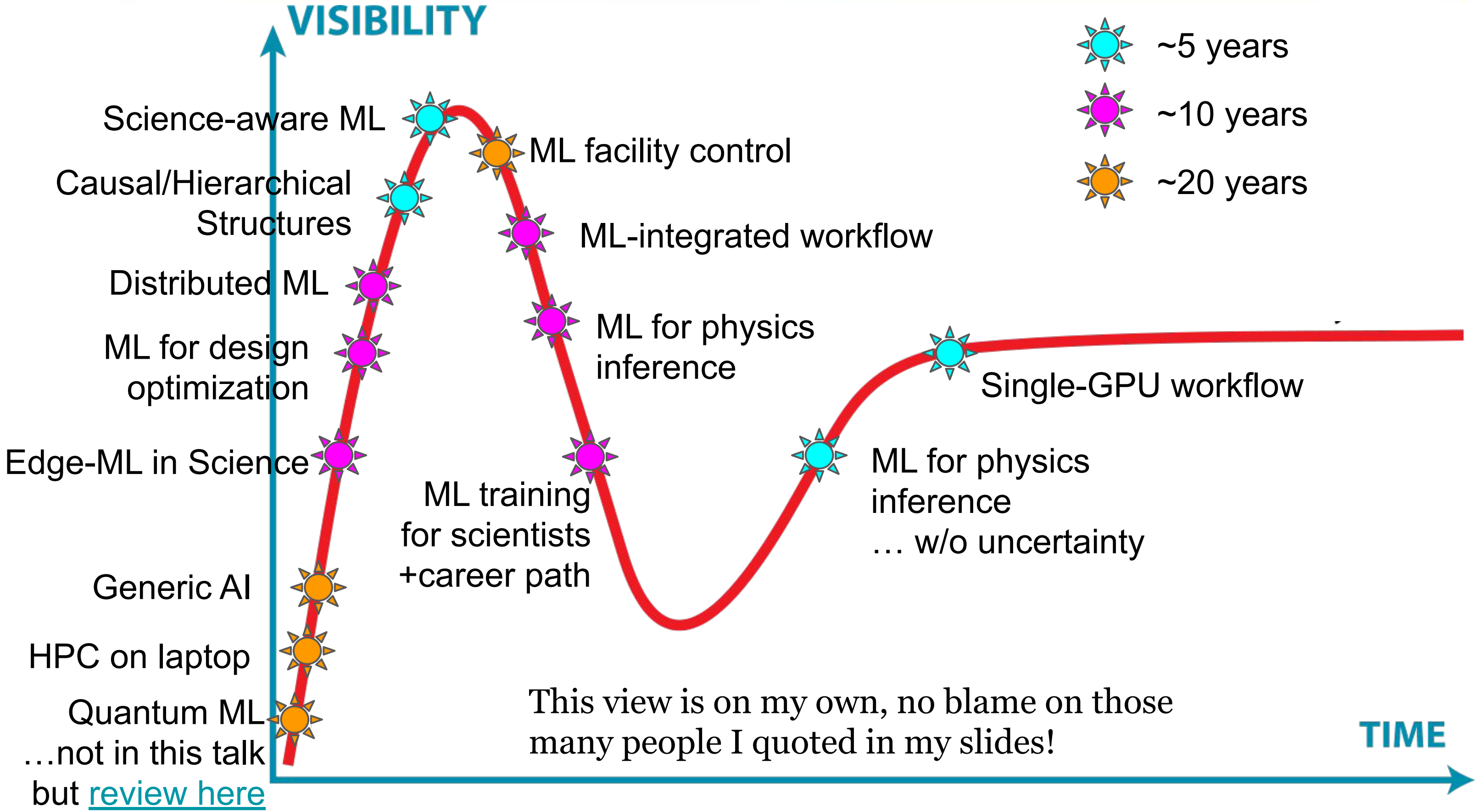
CMS and ATLAS Computing Trainings
Both had short sessions (3-4 hours) on ML, led by experts in the experiments (Experiment focused)



CoDaS-HEP School
5 day advanced computing training including ~2 days on ML led by domain experts with physics connections (HEP focused)

Slides by Savannah Thais (Snowmass ML group workshop)

Machine Learning Discoveries
So… Hype Cycle?



This view is on my own, no blame on those many people I quoted in my slides!

THANK YOU
for your attention!

Questions?
(do I have time?)

My daughter, who knows "20 years later"?

She'll be done being a teen-ager, though! ;)