Simulation session

Corey Adams (ANL), Makoto Asai (SLAC)

14:00 → 17:00 **Day 1 afternoon**

Conveners: Corey Adams (ANL) , Makoto Asai (SLAC)

14:00 **Simulations: Introduction** 5m

14:05 **Event Generation and Simulation Needs for the EIC** 25m

Speaker: Dr Markus Diefenthaler (Jefferson Lab)

Diefenthaler-AI4EIC...

14:30 **Bottlenecks and limitations in classical simulations: where can AI help?** 25m

Speaker: Sylvester Joosten (Argonne National Laboratory)

20210907-AI4EIC-S...

14:55 **What AI offers for EIC simulations** 30m

Speaker: Michelle Kuchera (Davidson College)

Kuchera-AI4EIC-ex...

15:25 **break** 15m

15:40 **Generative AI applications for simulations in colliders** 30m

Speaker: Benjamin Nachman

SimulationAI4EIC_...

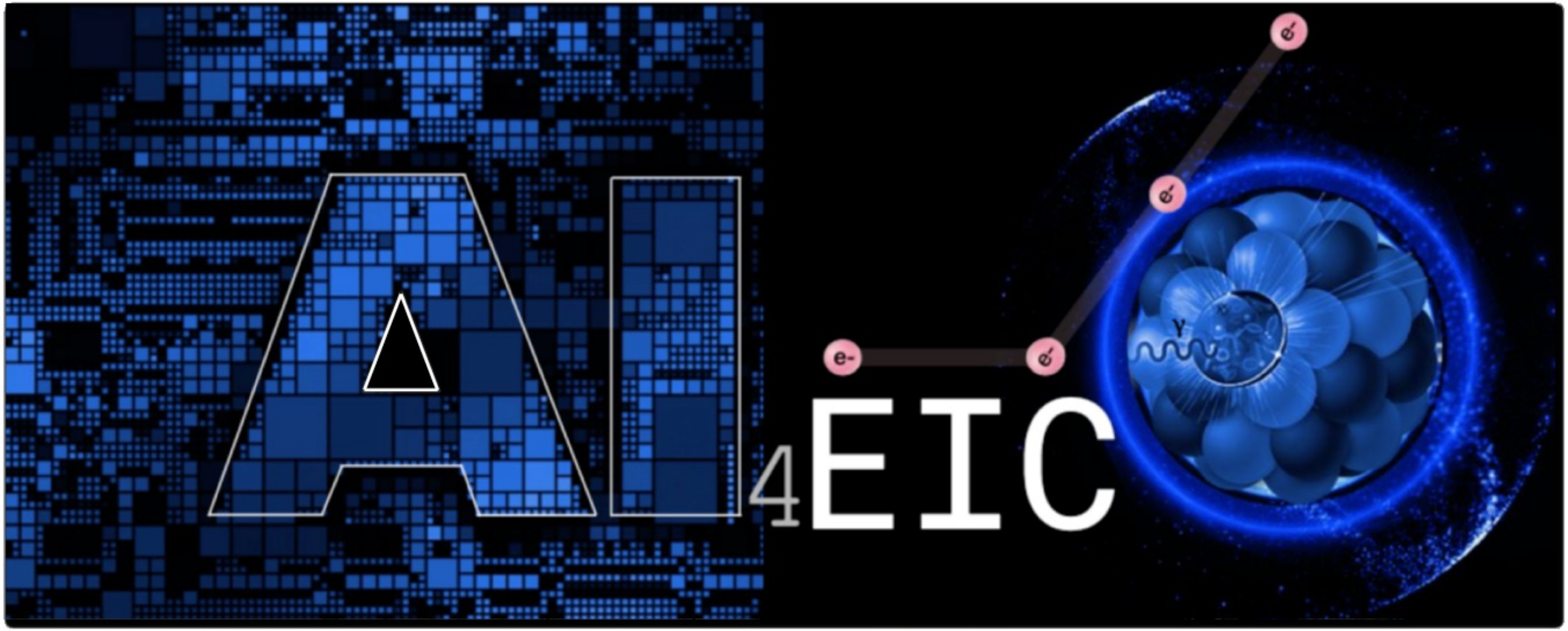
16:10 **Machine Learning for the LHCb simulations** 25m

Speaker: Lucio Anderlini (INFN)

Machine Learning f...

16:35 **Discussion** 25m

Event Generation and Simulation Needs for the EIC



Markus Diefenthaler

EIC²
EIC Center at Jefferson Lab

Jefferson Lab

The role of AI/ML in simulations

Lesson learned High-precision QCD measurements require high-precision simulations

Statistical accuracy for precise hypothesis testing
- up to trillion of simulated events required (HL-LHC)
- often computationally intensive, in particular calorimeter simulations

Common alternatives
- fast simulations with computationally efficient approximations, e.g., parameterizations or look-up tables
- **still** insufficient accuracy for high-precision measurements

Promising alternatives
- fast generative models, e.g., GANs or VAEs
- AI driven design, e.g., Bayesian optimization


Eur. Phys. J. A (2021) 57:100
https://doi.org/10.1140/epja/s10050-020-00290-x

THE EUROPEAN PHYSICAL JOURNAL A

Review

A.I. for nuclear physics

Paulo Bedaque[1], Amber Boehnlein[2,a], Mario Cromaz[3], Markus Diefenthaler[2], Latifa Elouadrhiri[2], Tanja Horn[4], Michelle Kuchera[5], David Lawrence[2], Dean Lee[6], Steven Lidia[6], Robert McKeown[2], Wally Melnitchouk[2], Witold Nazarewicz[6], Kostas Orginos[2,7], Yves Roblin[2], Michael Scott Smith[8], Malachi Schram[9], Xin-Nian Wang[3]

[1] University of Maryland, College Park, MD, USA
[2] Thomas Jefferson National Accelerator Facility, Newport News, VA, USA
[3] Lawrence Berkeley National Laboratory, Berkeley, CA, USA
[4] Catholic University, Washington D.C., USA
[5] Davidson College, Davidson, NC, USA
[6] Michigan State University, East Lansing, MI, USA
[7] College of William & Mary, Williamsburg, VA, USA
[8] Oak Ridge National Laboratory, Oak Ridge, TN, USA
[9] Pacific Northwest National Laboratory, Richland, WA, USA

Received: 9 September 2020 / Accepted: 7 October 2020 / Published online: 22 March 2021

Communicated by Ulf Meissner

Abstract This report is an outcome of the workshop AI for Nuclear Physics held at Thomas Jefferson National Accelerator Facility on March 4–6, 2020

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics under contract DE-AC05-06OR23177. Participation of students and early career professionals was supported by NSF, Division of Physics, under the Grant 'Artificial Intelligence (AI) Workshop in Nuclear Physics,' Award Number 2017170. Support for the Hackathon was provided by the University of Virginia School of Data Sciences and by Amazon Web Services.

This report is an outcome of the workshop *AI for Nuclear Physics* held at Thomas Jefferson National Accelerator Facility on March 4–6, 2020. The workshop brought together 184 scientists to explore opportunities for Nuclear Physics in the area of Artificial Intelligence. The workshop consisted of plenary talks, as well as six working groups.

[a] e-mail: amber@jlab.org (corresponding author)

Contents

Springer

Jefferson Lab

Markus Diefenthaler

mdiefent@jlab.org

Summary

- **Simulations** essential for design of experiments, data analysis, and verification of measurements.
- **Simulations** for the **EIC**, i.e. MCEGs and fast and full detector simulations for the EIC, require **R&D**. We miss core capabilities and we need to work towards accuracy and precision.
- **Simulation R&D** is most efficiently done in common projects and in collaboration with other fields, e.g., HEP or data science.
- Many opportunities for AI/ML to complement and improve **simulations**. While AI/ML approaches will substitute part of **simulation workflows**, they will not replace core tools, e.g., general-purpose MCEGs or Geant4.

This work is supported by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics, under contract DE-AC02-06CH11357

Argonne NATIONAL LABORATORY | 75 1946–2021

BOTTLENECKS IN CLASSICAL SIMULATIONS: WHERE CAN AI HELP?

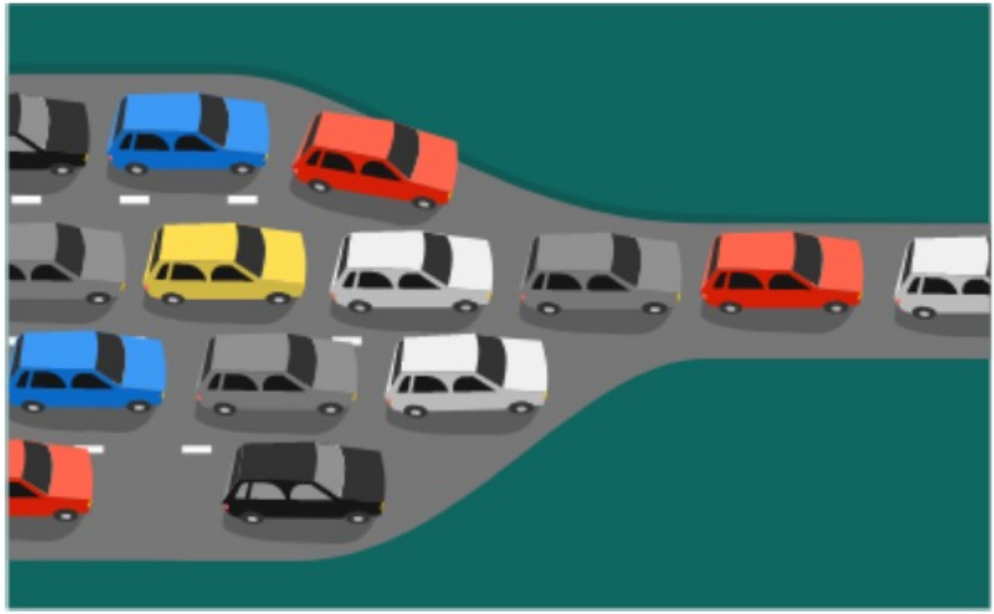
SYLVESTER JOOSTEN
sjoosten@anl.gov

AI4EIC Workshop
Tuesday, September 7, 2021

SIMULATION BOTTLENECKS

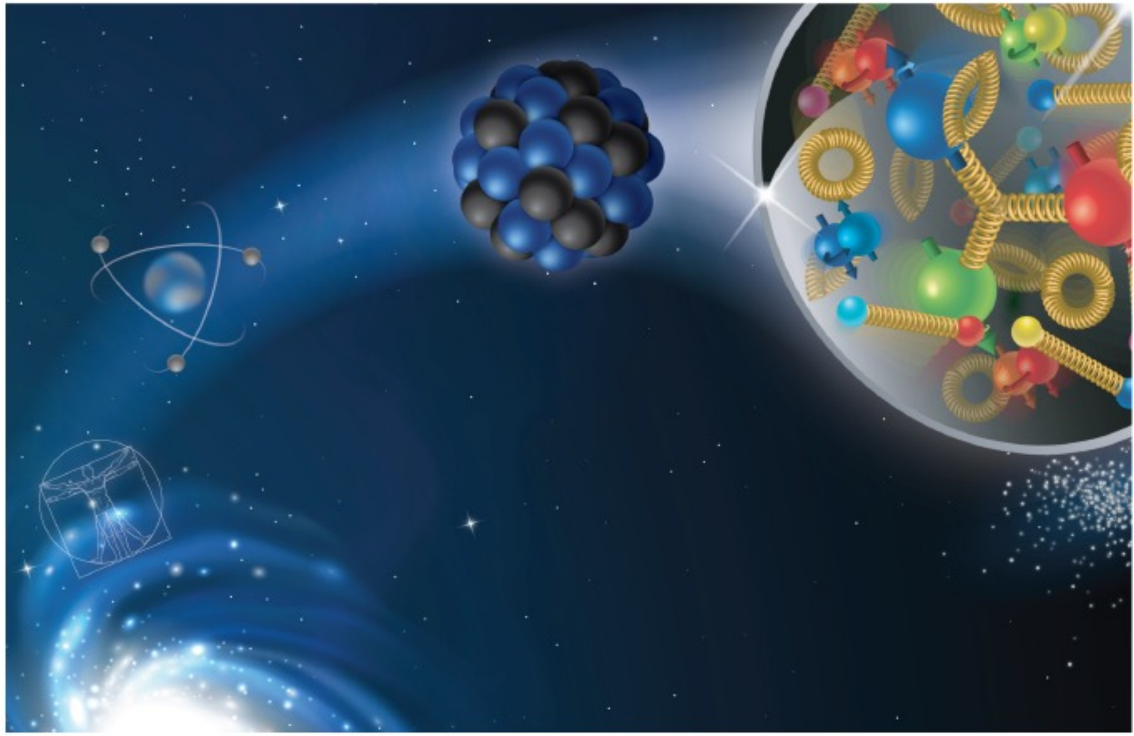
Many particles, many components, many steps

- Usually bottlenecks occur where the particle count is high, e.g. as part of a calorimeter shower, or optical photons in a RICH.
- Bottlenecks can also occur in when navigating very detailed geometries (e.g. fiber calorimeters with millions of fibers).
- Finally, scenarios where we need many precise steps through a magnetic field (for upstream & downstream near-beamline detection) is relatively expensive.
- Often multiple bottlenecks at once.



SUMMARY

- Short-term: need large-scale simulations for optimization of a complex detector system.
- Long-term: need (even larger)-scale simulations to properly analyze high-luminosity/high-precision measurements.
- Bottlenecks usually a combination between many particles, many geometry elements and/or many simulation steps.
- Calorimetry, Cherenkov detectors and the far-forward/far-backward regions are prime targets for AI-driven acceleration.



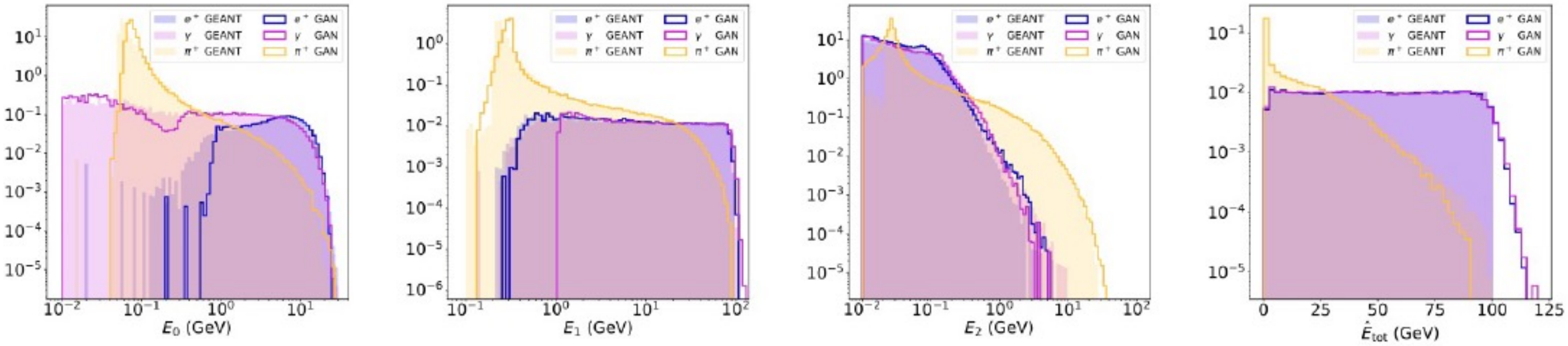
What can AI offer for Simulation at the EIC?

Michelle P. Kuchera | 7 September 2021

DAVIDSON

Generative Adversarial Networks

Simulation



caloGAN (ATLAS)
Paganini, et. al [1712.10321]

Silicon- Tungsten calorimeter of the proposed International Large Detector
Buhmann , et. al [2005.05334]

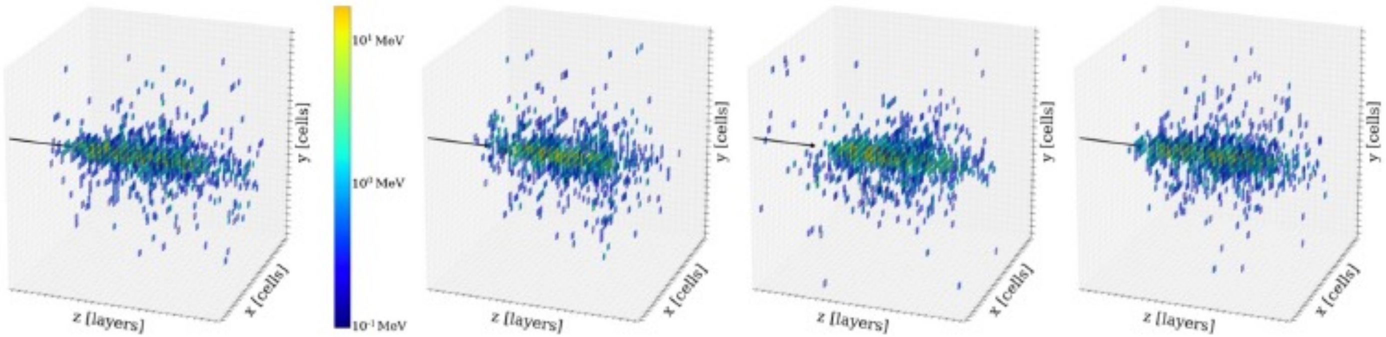
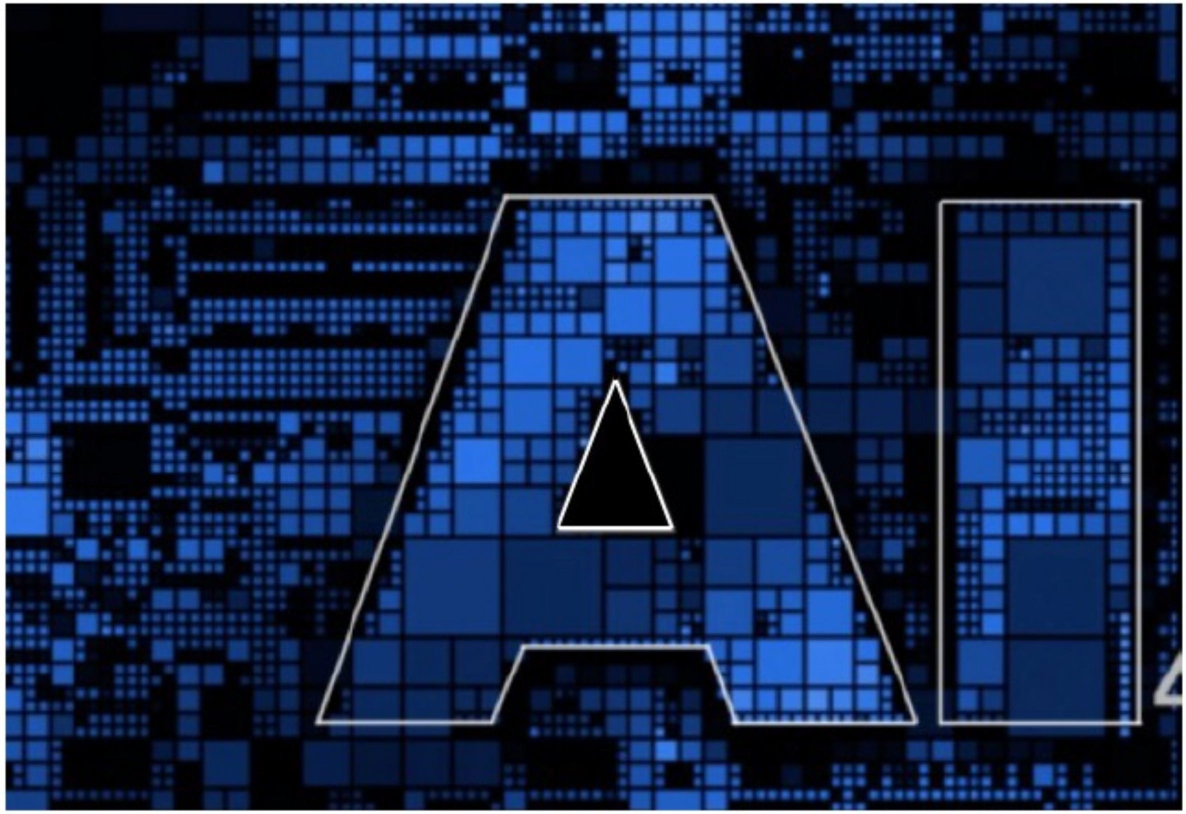


Fig. 5 Examples of individual 50 GeV photon showers generated by Geant4 (left), the GAN (center left), WGAN (center right), and BIB-AE (right) architectures. Colors encode the deposited energy per cell.

Michelle Kuchera
mikuchera@davidson.edu

Summary

- **Ai has the potential for large impacts on Simulation for the EIC.**
- **Large body** of prior related work. Often at "bleeding edge" of AI research. Less commonly used for simulation in practice. Requires work.
- **Simulation R&D** is most efficiently done in common projects and in collaboration with other fields, e.g., HEP or data science.
- **Do not expect replacement** of core tools, e.g., general-purpose MCEGs or Geant4.



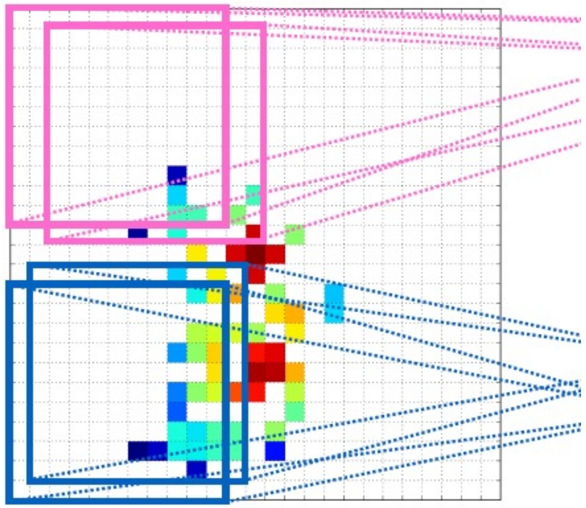
Generative ML applications for simulations in colliders

Benjamin Nachman

Lawrence Berkeley National Laboratory

bpnachman.com
bpnachman@lbl.gov

@bpnachman bnachman



BERKELEY LAB

BIDS
BERKELEY INSTITUTE FOR DATA SCIENCE

AI4EIC
September 7, 2021

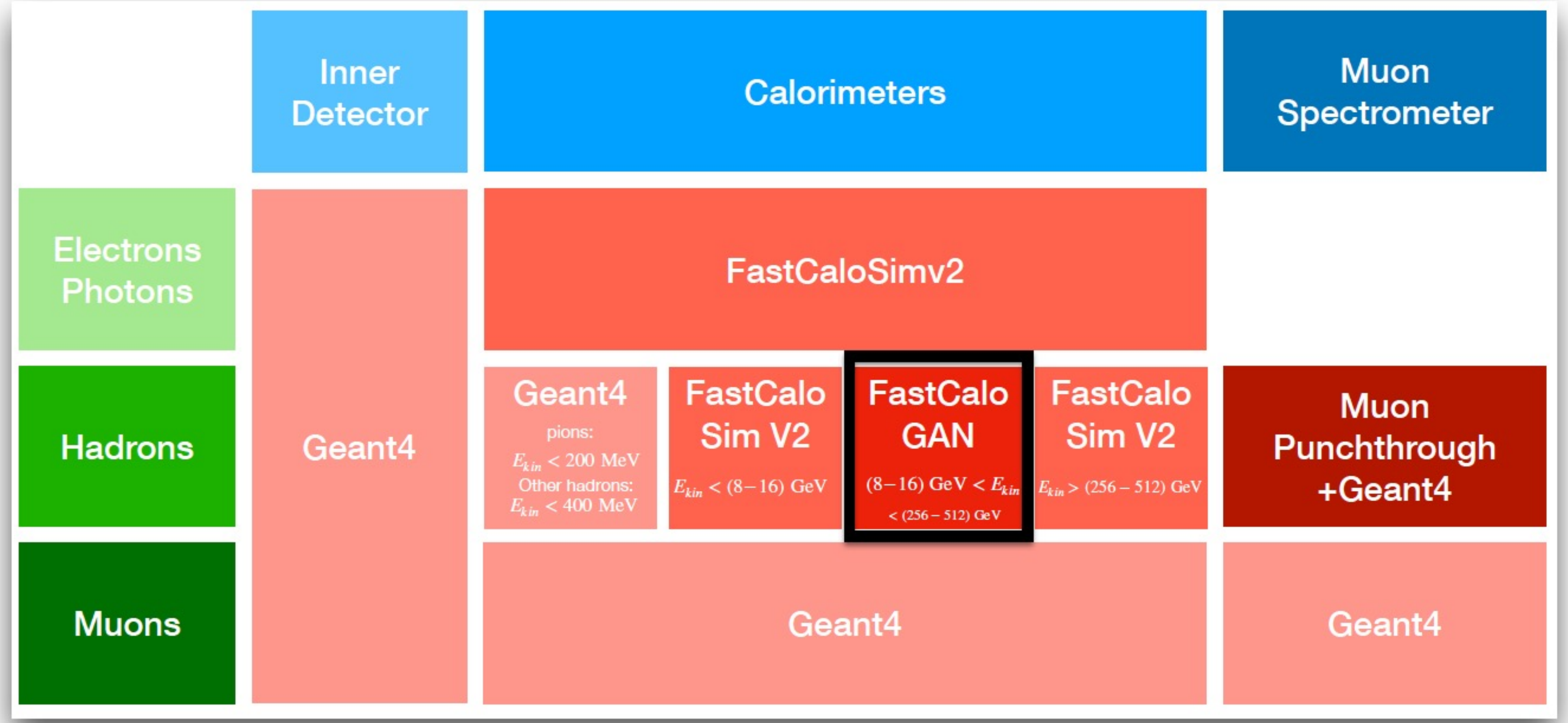
M. Paganini, L. de Oliveira, B. Nachman, 1705.02355

Timing

Generation Method	Hardware	Batch Size	milliseconds/shower
GEANT4	CPU	N/A	1772
CALOGAN	CPU *Intel Xeon E5-2670*	1	13.1
		10	5.11
		128	2.19
		1024	2.03
	GPU *NVIDIA K80*	1	14.5
		4	3.68
		128	0.021
		512	0.014
		1024	0.012

(clearly these numbers have changed as both technologies have improved - this is simply meant to be qualitative & motivating!)

Integration into real detector sim.



The ATLAS Collaboration fast simulation (AF3) now includes a GAN at intermediate energies for pions

Conclusions and Outlook

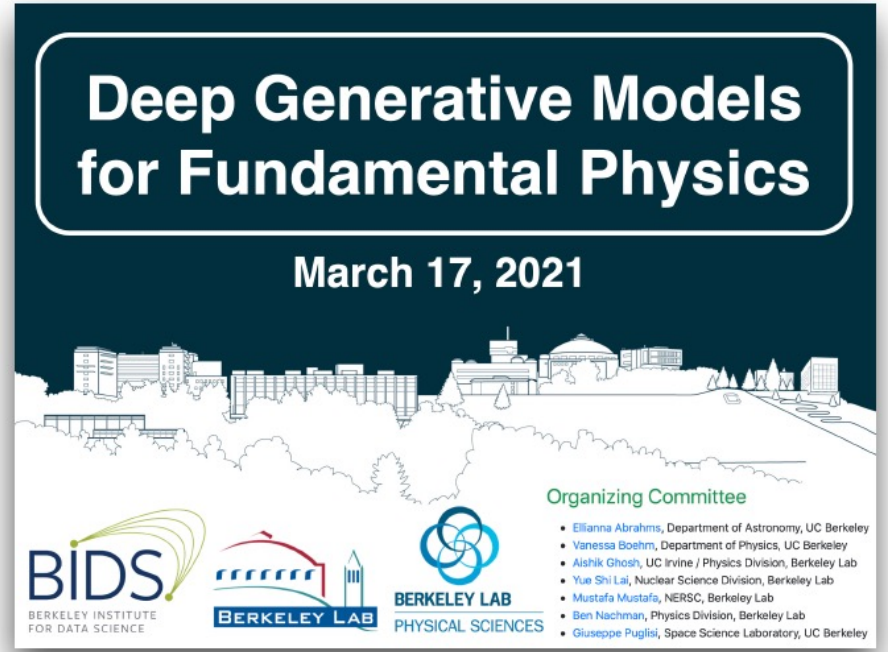
Generative models hold great promise for enhancing, supplanting, and extending simulations for collider physics

The examples I gave today were not comprehensive

- see the Living Review for more references

All of the techniques I discussed today could be used for **physics at the EIC**!

This is a link to a recent Berkeley workshop dedicated to generative models →

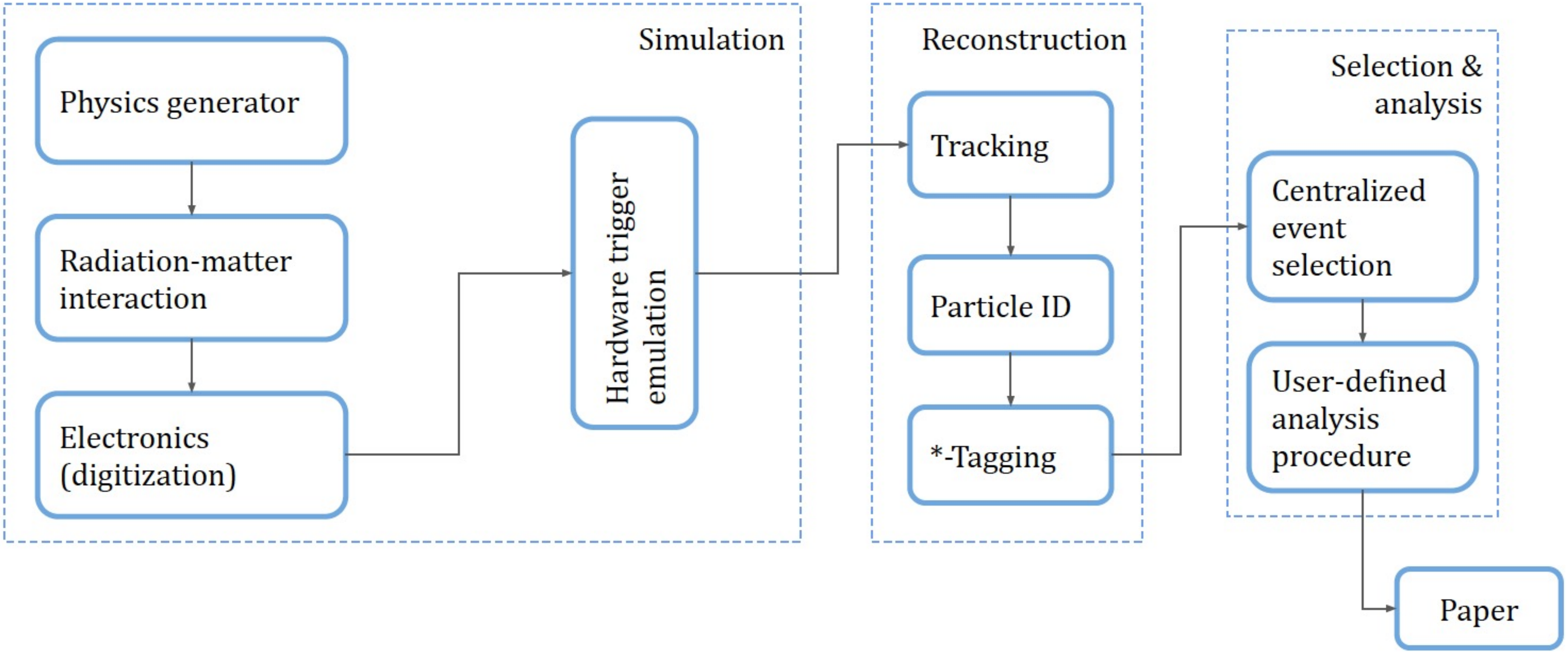


Machine Learning for the LHCb Simulations

Lucio Anderlini on behalf of the LHCb Collaboration
Istituto Nazionale di Fisica Nucleare – Sezione di Firenze

2021 - 09 - 07

Standard simulation: the big picture



Our ML building block

$$\underbrace{x}_{\text{Target variable}} = F^{-1}(\underbrace{u}_{\text{Random noise}};\underbrace{p,\eta,\ldots}_{\text{Conditions}})$$

Random noise

Conditions (e.g. momentum)

Generative Adversarial Networks

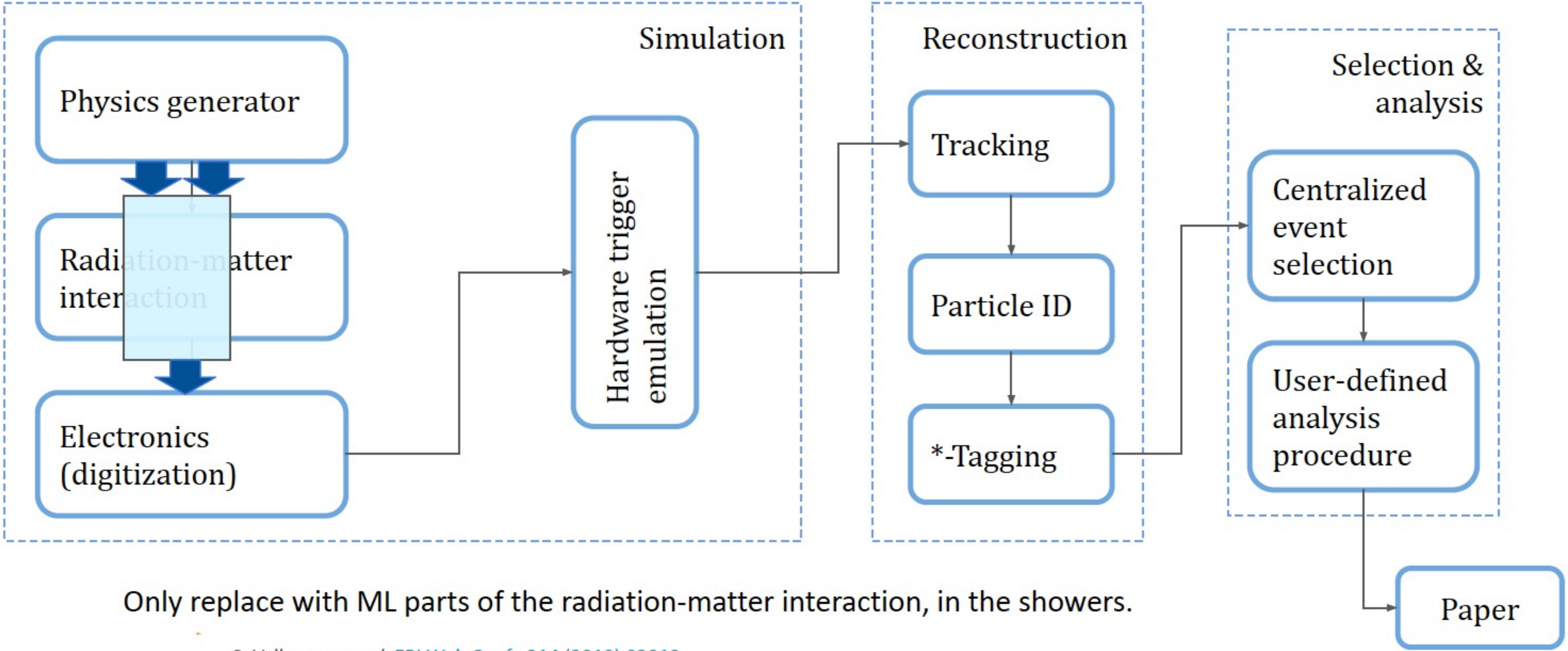
Normalizing Flows

Variational AutoEncoder

…

Some detector-related **quantities**

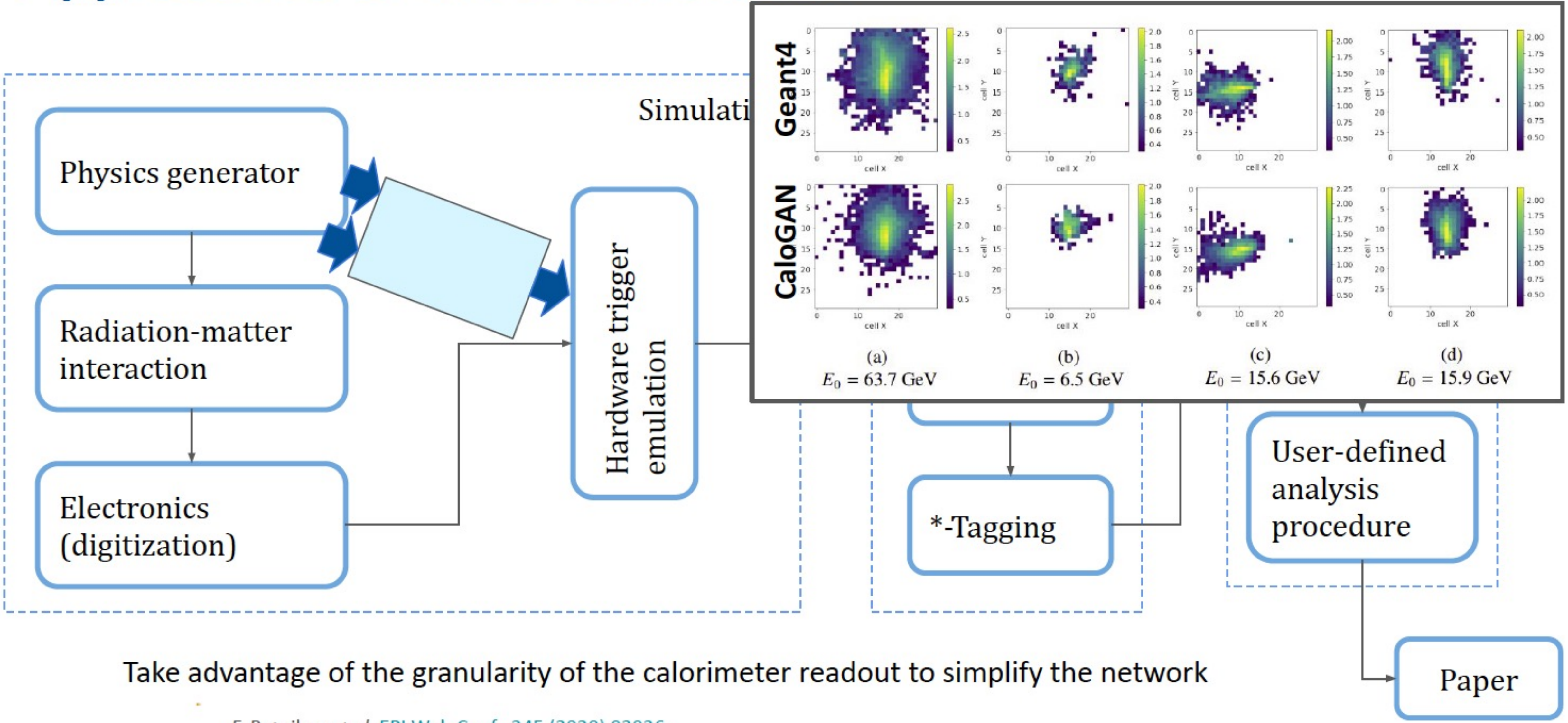
Approaches to ML in simulation



Only replace with ML parts of the radiation-matter interaction, in the showers.

S. Vallecorsa *et al.* EPJ Web Conf., 214 (2019) 02010

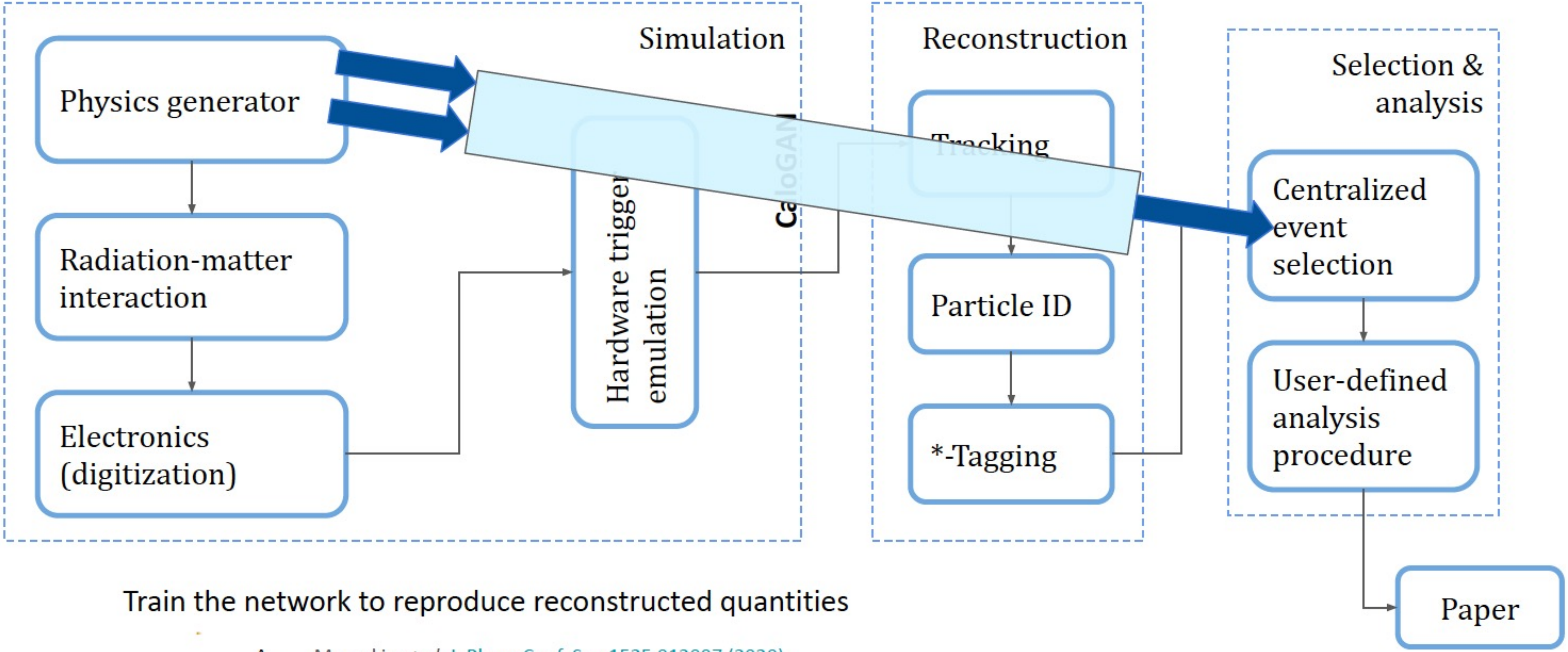
Approaches to ML in simulation



Take advantage of the granularity of the calorimeter readout to simplify the network

F. Ratnikov *et al.* [EPJ Web Conf., 245 (2020) 02026]

Approaches to ML in simulation



Train the network to reproduce reconstructed quantities

A. Maevskiy *et al.* J. Phys.: Conf. Ser. 1525 012097 (2020)

Conclusion 2

Many studies are ongoing to use machine learning to speed up the simulation, taking different approaches:

- simulating the radiation-matter interaction faster
- simulating the response of some detector
- simulating the whole simulation pipeline to reconstructed analysis level quantities

The various solutions are all important as they can all help speeding the simulation up.

To sum up

- Simulation is one of the least advanced areas in the use of AI/ML.
    - Many ideas, opportunities
    - Require works – resource allocation needed
- Two phases
    - Simulation-driven detector optimization with boost of AI/ML (~ 5 years)
    - Fully tailored simulation with maximum use of AI/ML to meet the statistics requirements (~ decade(s))
- Do not expect AI/ML to fully replace core tools, e.g. MCEGs or Geant4.
    - Training, validation
- Modular approach
    - Calorimeter, optical photon detector for PID, very-forward detectors, machine-detector interface, background (neutron, beam halo, etc.) are good start.
- Simulation R&D is most efficiently done in common project and in collaboration with other fields. Engagement with broader AI/Comp.Sci. community is valuable
    - Expect engagement to include data sharing and bidirectional contributions