Machine Learning at LHC : from state of the art to future developments

Sanmay Ganguly
IIT-Kanpur

02/02/2024

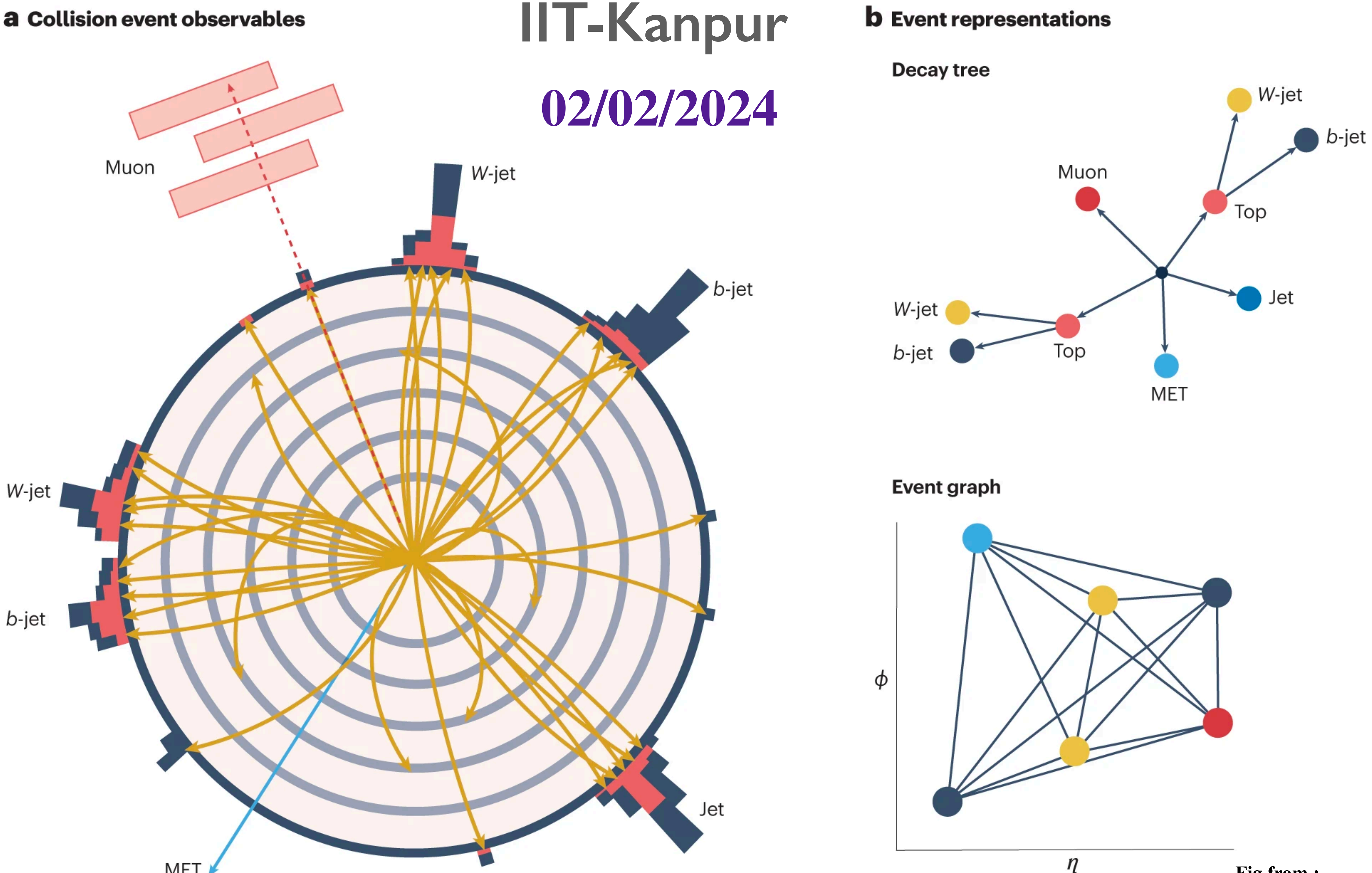


Fig from :
Nat Rev Phys 5, 281–303 (2023)

Outline

- Introduction
- Current HEP-ML activities
- What's brewing in the community?
- Activities in broader HEP community
- Summary

ML and HEP are old friends

TRACK FINDING WITH NEURAL NETWORKS

Nuclear Instruments and Methods in Physics Research A279 (1989) 537–545
North-Holland, Amsterdam

Carsten PETERSON

Department of Theoretical Physics, University of Lund, Sölvegatan 14A, S-22362 Lund, Sweden

Received 6 December 1988 and in revised form 6 March 1989

$$V_{ij} = \tanh\left[\left(\sum_{l} \frac{\cos^m \theta_{ijl}}{r_{ij} + r_{jl}} - \alpha\left(\sum_{l \neq j} V_{il} + \sum_{k \neq i} V_{kj}\right) - \beta\left(\sum_{k \neq l} S_{kl} - N\right)\right)/T\right], \tag{24}$$

C. Peterson / Track finding with neural networks

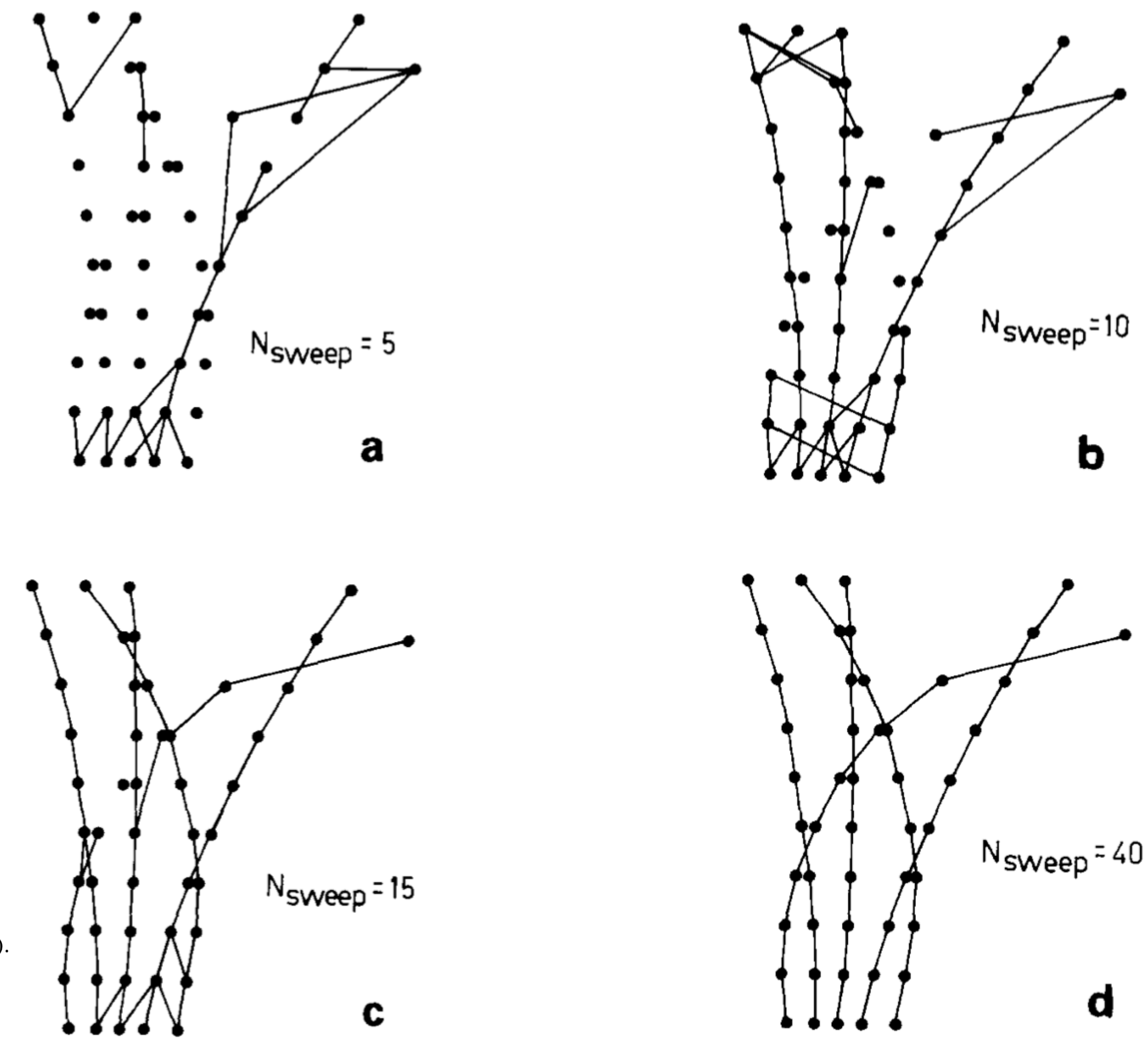


Fig. 9. Segments with $V_{ij} > 0.1$ at different evolution stages of the MFT equations.

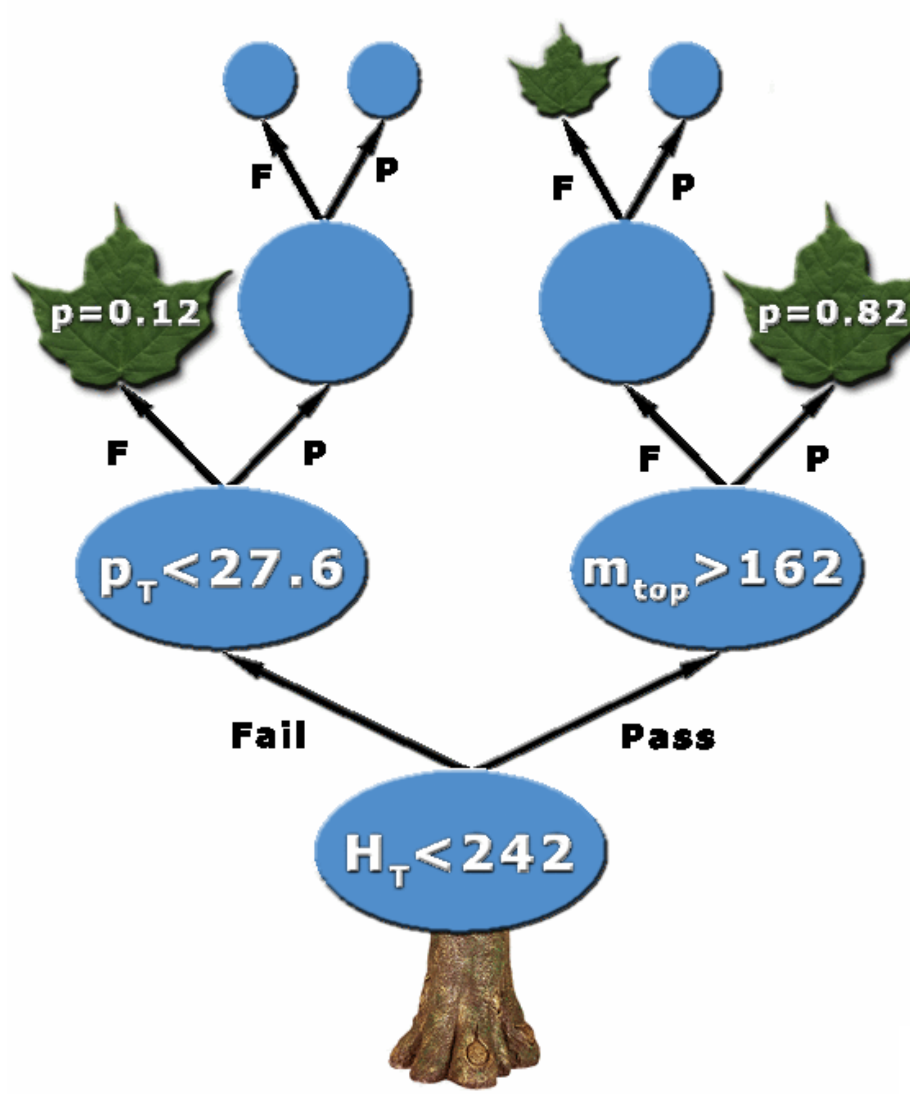
[1] L. Lönnblad, C. Peterson and T. Rögnvaldsson, ``Pattern Recognition in High Energy Physics with Artificial Neural Networks", *Comput. Phys. Commun.* **70**, 167 (1992).

[2] L. Lönnblad, C. Peterson and T. Rögnvaldsson, "Using Neural Networks to Identify Jets", *Nucl. Phys.* **B 349**, 675 (1991).

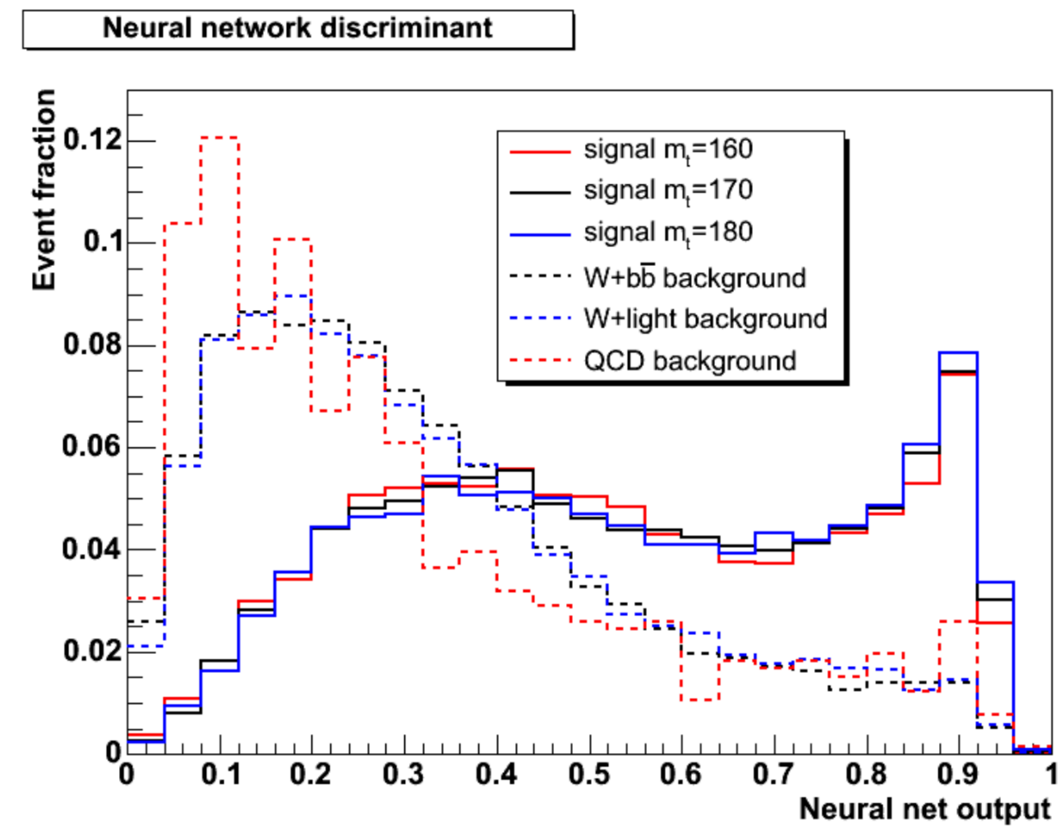
[3] L. Lönnblad, C. Peterson and T. Rögnvaldsson, "Finding Gluon Jets with a Neural trigger", *Phys. Rev. Lett.* **65**, 1321 (1990).

[4] L. Lönnblad, C. Peterson, H. Pi and T. Rögnvaldsson, ``Self-organizing Networks for Extracting Jet Features", *Comput. Phys. Commun.* **67**, 193 (1991).

ML and HEP are old friends



Boosted decision trees were used in early 90's in Tevatron to discover top quarks



Machine learning for event selection in high energy physics

Shimon Whiteson [a,*], Daniel Whiteson [b]

[a] *Informatics Institute, University of Amsterdam, Science Park 107, 1098 XG Amsterdam, The Netherlands*
[b] *Department of Physics and Astronomy, University of California, Irvine, 4129 Frederick Reines Hall, Irvine, CA 92697-4575, USA*

The use of neural networks were ventured from early LHC days.

ARTICLE INFO

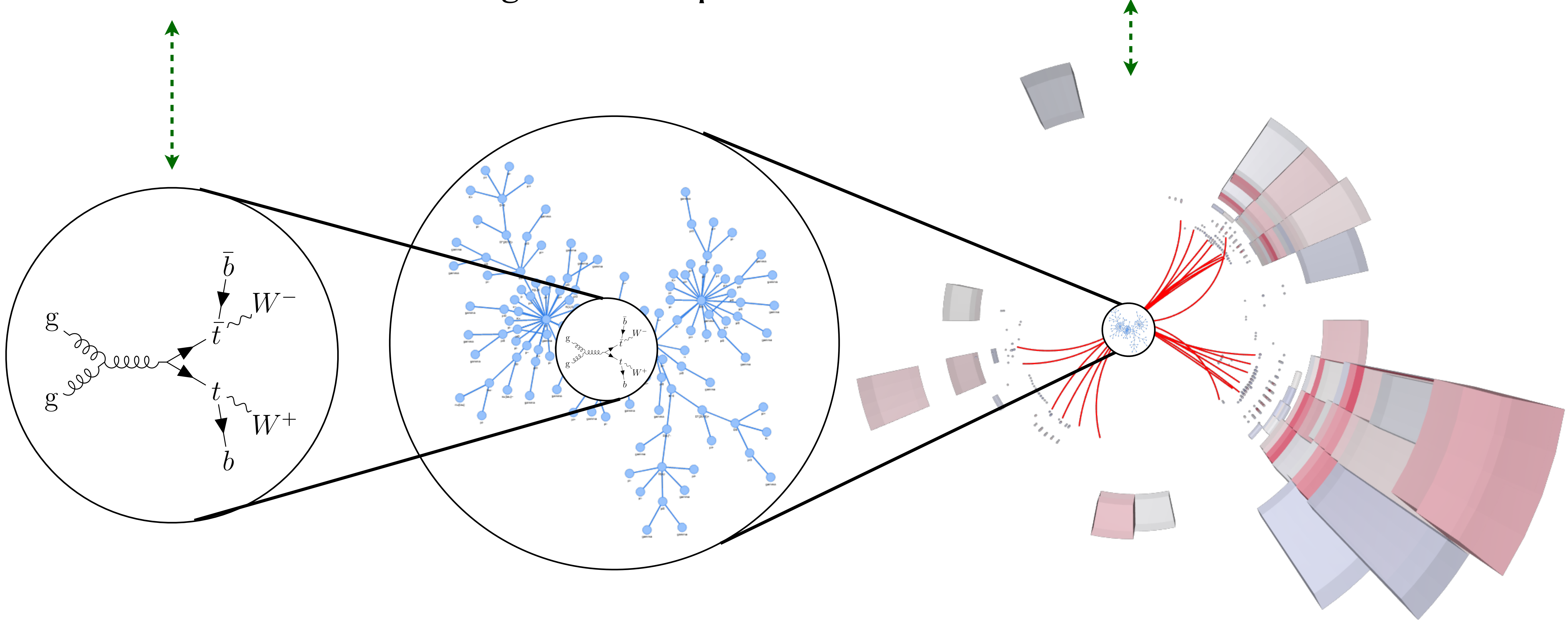
Article history:
Received 18 August 2008
Received in revised form
27 April 2009
Accepted 7 May 2009
Available online 13 June 2009

ABSTRACT

The field of high energy physics aims to discover the underlying stru
studying exotic particles, such as the top quark and Higgs boson
accelerators. Since such accelerators are extraordinarily expensiv
from the resulting data is essential. However, most accelerator
interest, so making effective measurements requires event sele
particles of interest (signal) are separated from events producing

A broad strategy towards physics inference

Guess the Lagrangian ← solving the inverse problem via ML — $\frac{1}{\mathscr{L}}\frac{d^2N}{dp_T d\eta}$



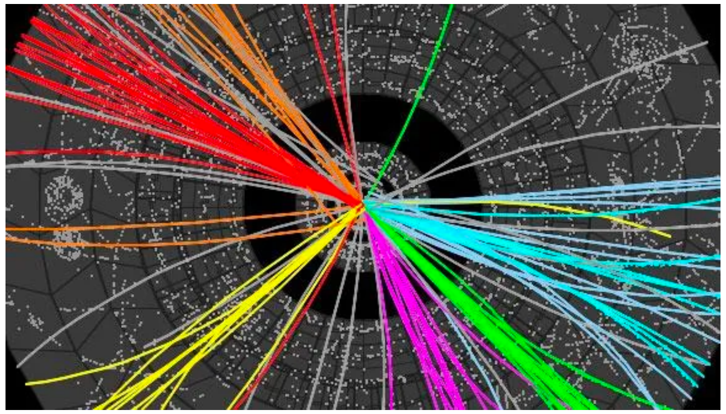
The physics at the core:
driven by the interaction between quantum fields, computed via perturbative or lattice techniques.

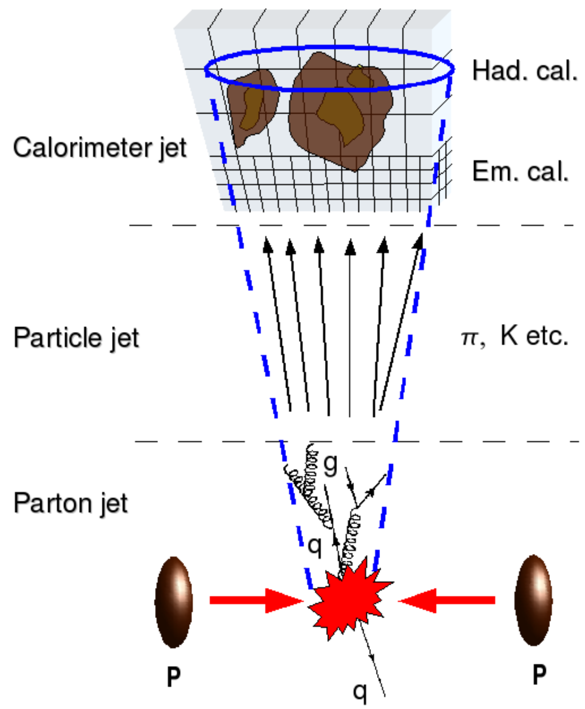
Final Particles
Produced via hadronization (no first principle analytic techniques are available)

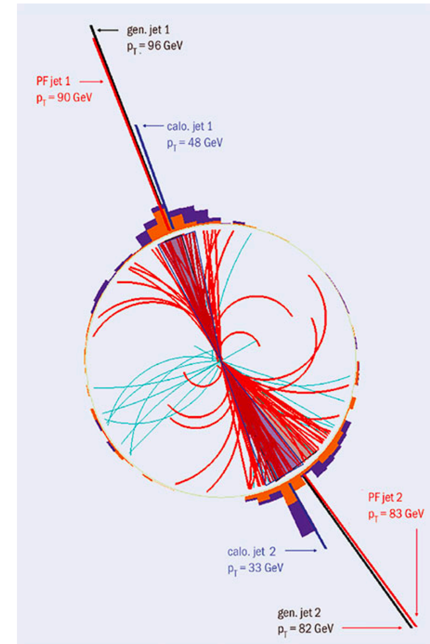
Detector output/Readout
Produced via hardware or simulation

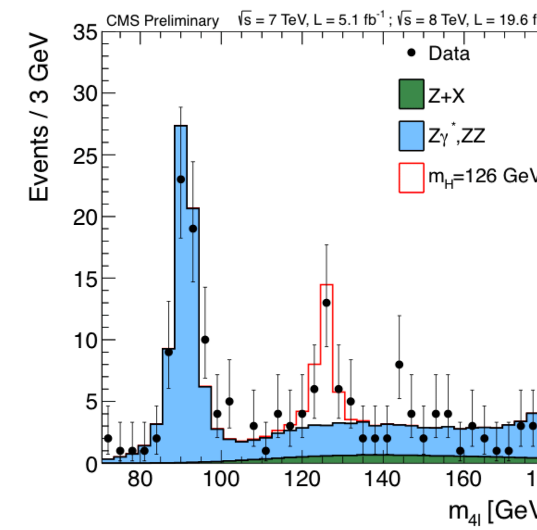
The LHC data flow-chain

ML can play a role at every instance of this flow chain.

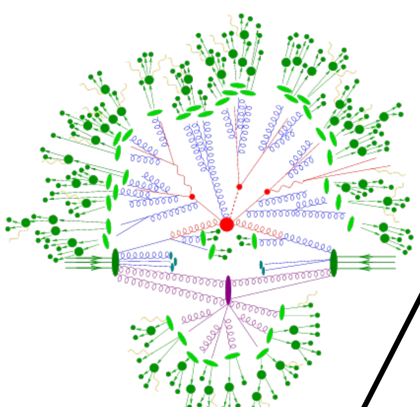


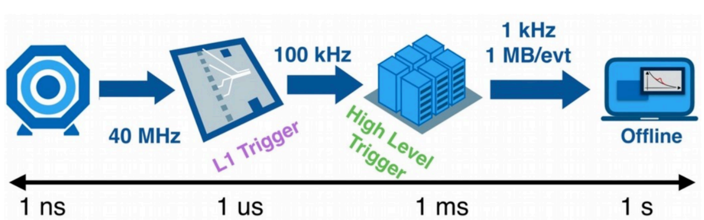


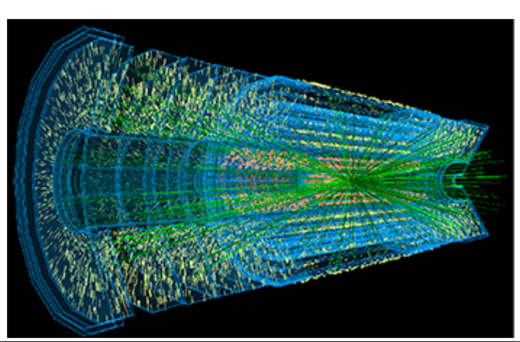


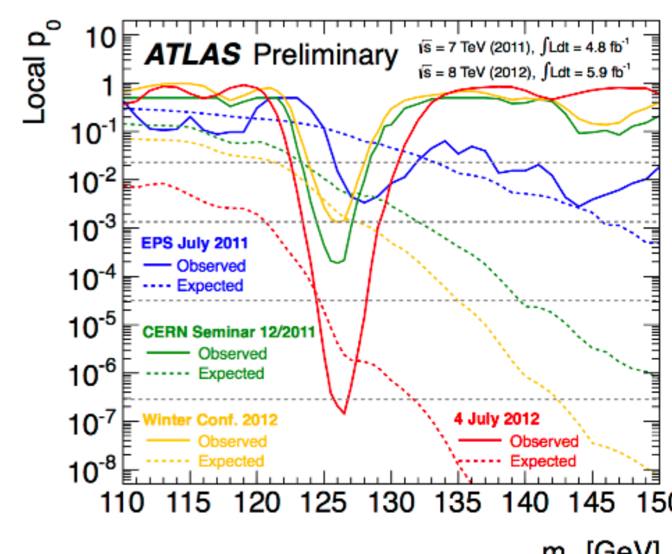


Trigger → Object rec → Event rec → Analysis → Inference





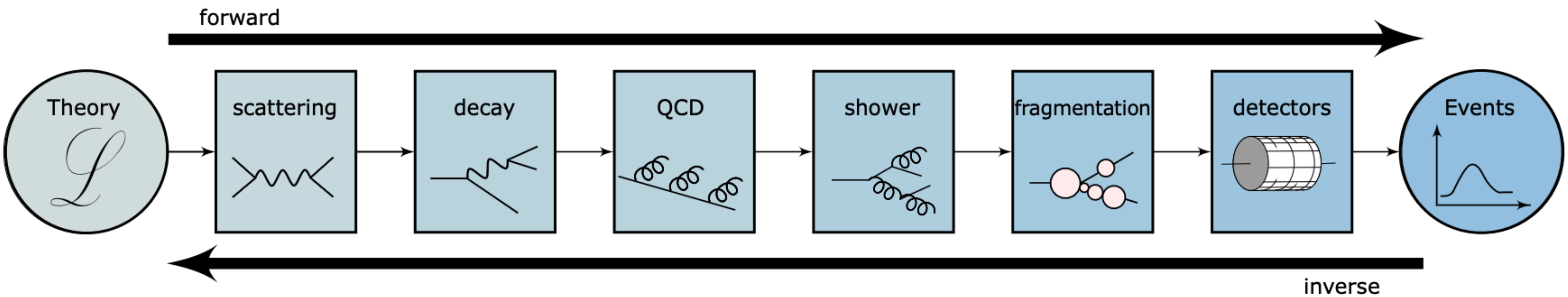




Inference

ME + PS generation
Detector Simulation

Monte-Carlo modelling using ML



arXiv > hep-ph > arXiv:2203.07460

Search...

Help | Advance

High Energy Physics – Phenomenology

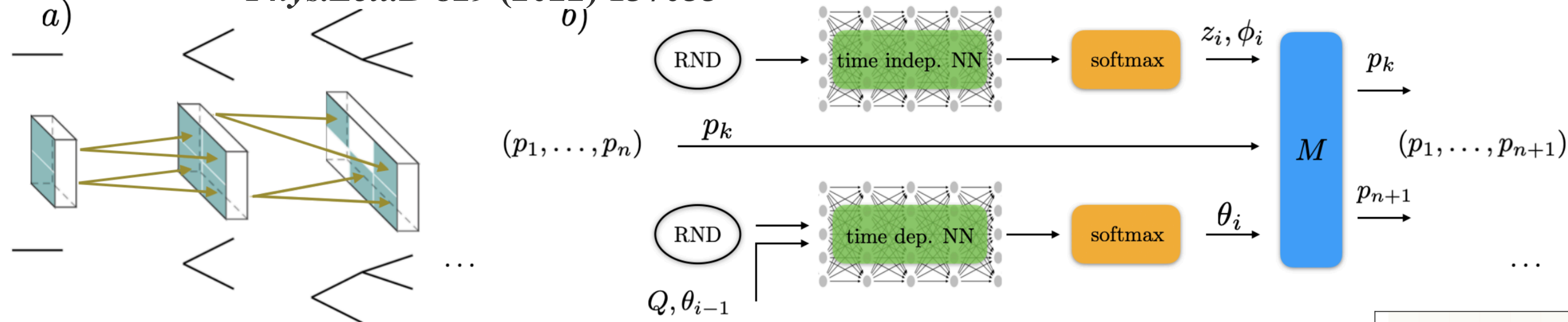
[Submitted on 14 Mar 2022 (v1), last revised 28 Dec 2022 (this version, v2)]

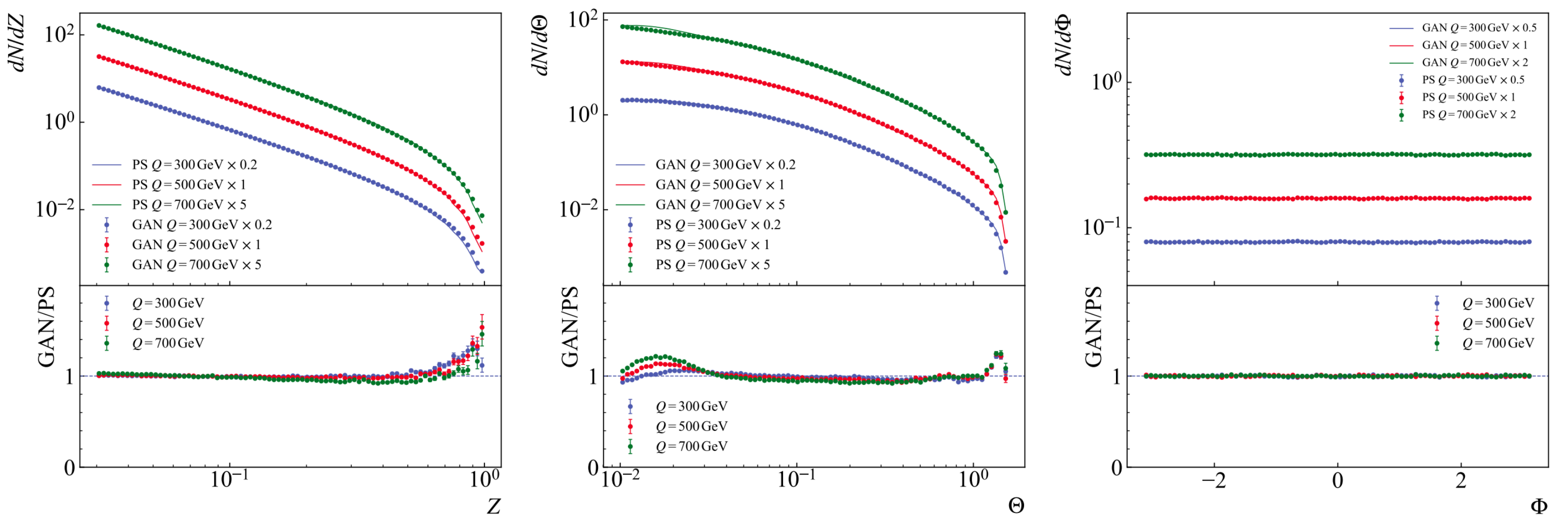
Machine Learning and LHC Event Generation

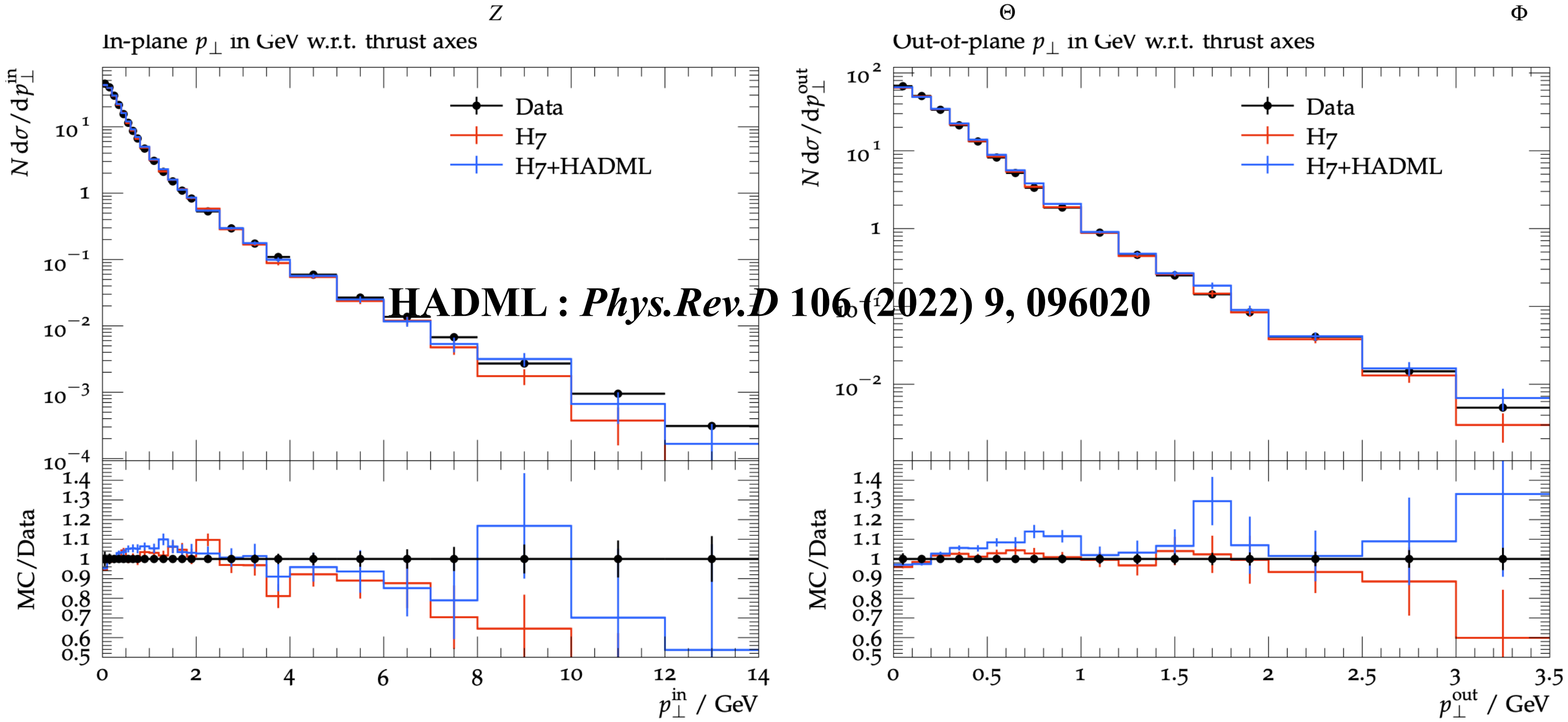
Anja Butter (ed), Tilman Plehn (ed), Steffen Schumann (ed), Simon Badger, Sascha Caron, Kyle Cranmer, Francesco Armando Di Bello, Etienne Dreyer, Stefano Forte, Sanmay Ganguly, Dorival Gonçalves, Eilam Gross, Theo Heimel, Gudrun Heinrich, Lukas Heinrich, Alexander Held, Stefan Höche, Jessica N. Howard, Philip Ilten, Joshua Isaacson, Timo Janßen, Stephen Jones, Marumi Kado, Michael Kagan, Gregor Kasieczka, Felix Kling, Sabine Kraml, Claudius Krause, Frank Krauss, Kevin Kröninger, Rahool Kumar Barman, Michel Luchmann, Vitaly Magerya, Daniel Maitre, Bogdan Malaescu, Fabio Maltoni, Till Martini, Olivier Mattelaer, Benjamin Nachman, Sebastian Pitz, Juan Rojo, Matthew Schwartz, David Shih, Frank Siegert, Roy Stegeman, Bob Stienen, Jesse Thaler, Rob Verheyen, Daniel Whiteson, Ramon Winterhalder, Jure Zupan

PS + Hadronization with ML

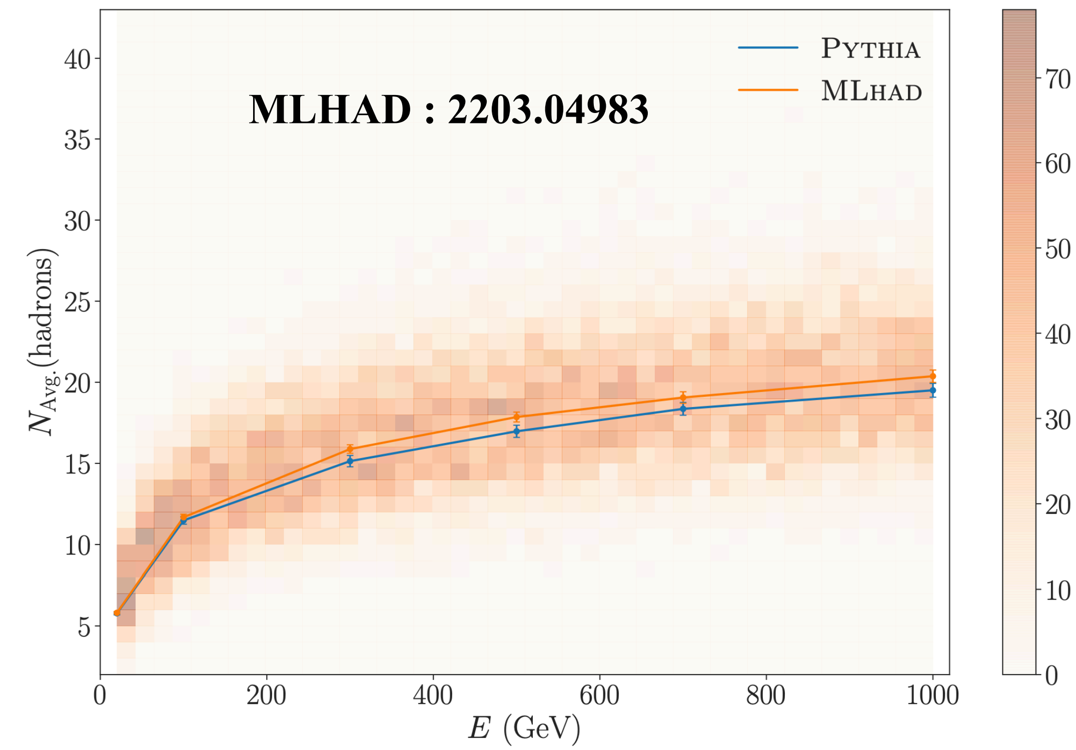
Phys.Lett.B **829 (2022) 137055**







HADML : ***Phys.Rev.D*** **106 (2022) 9, 096020**

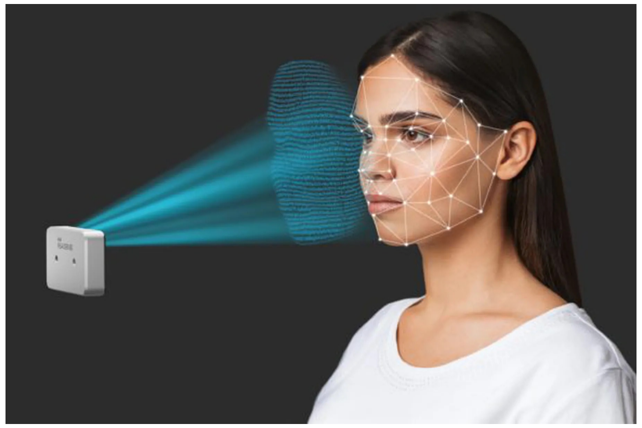


Overall shape normalization from Hadronization can be generated by generative models.

Energy cluster as images & a revolution begins

The energy deposition pattern can be thought as multi-layer image.

Image based methods were first to set in.



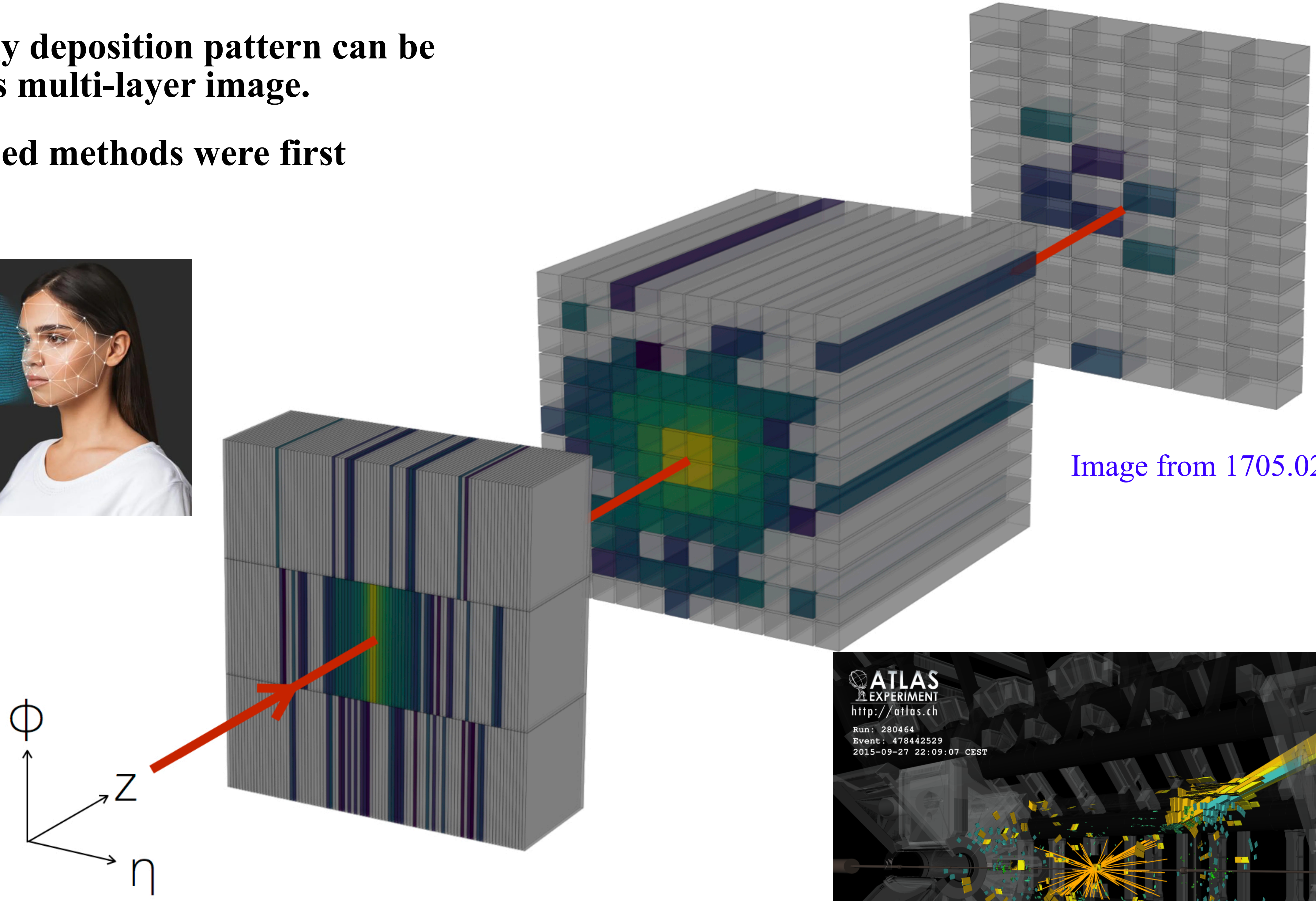
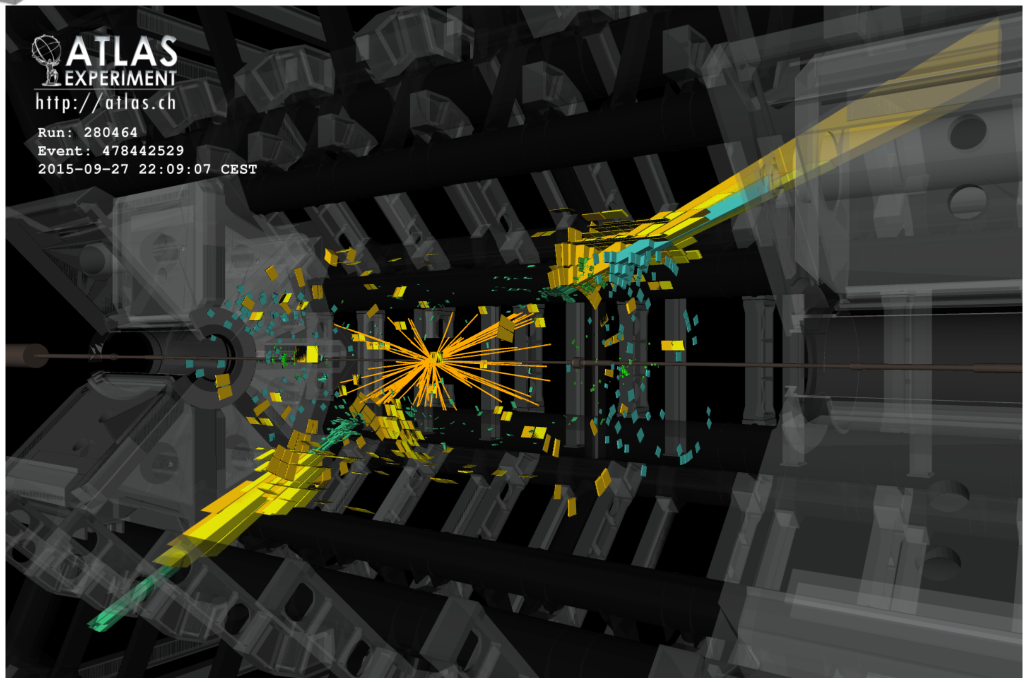
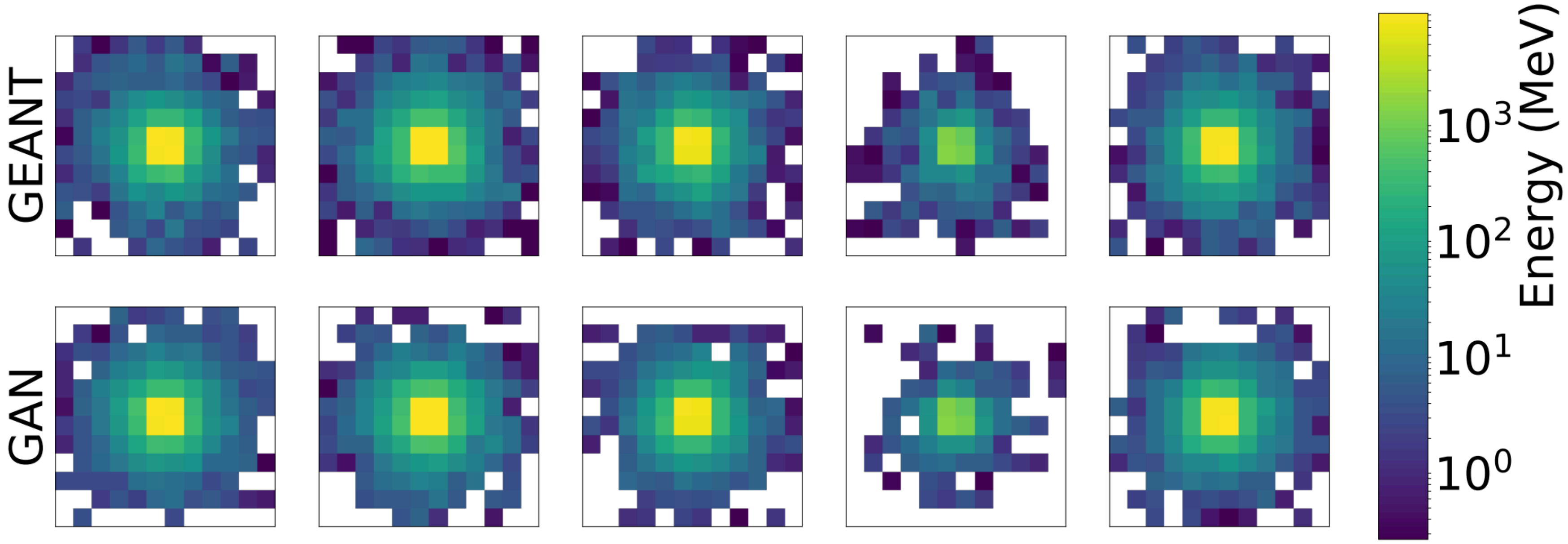


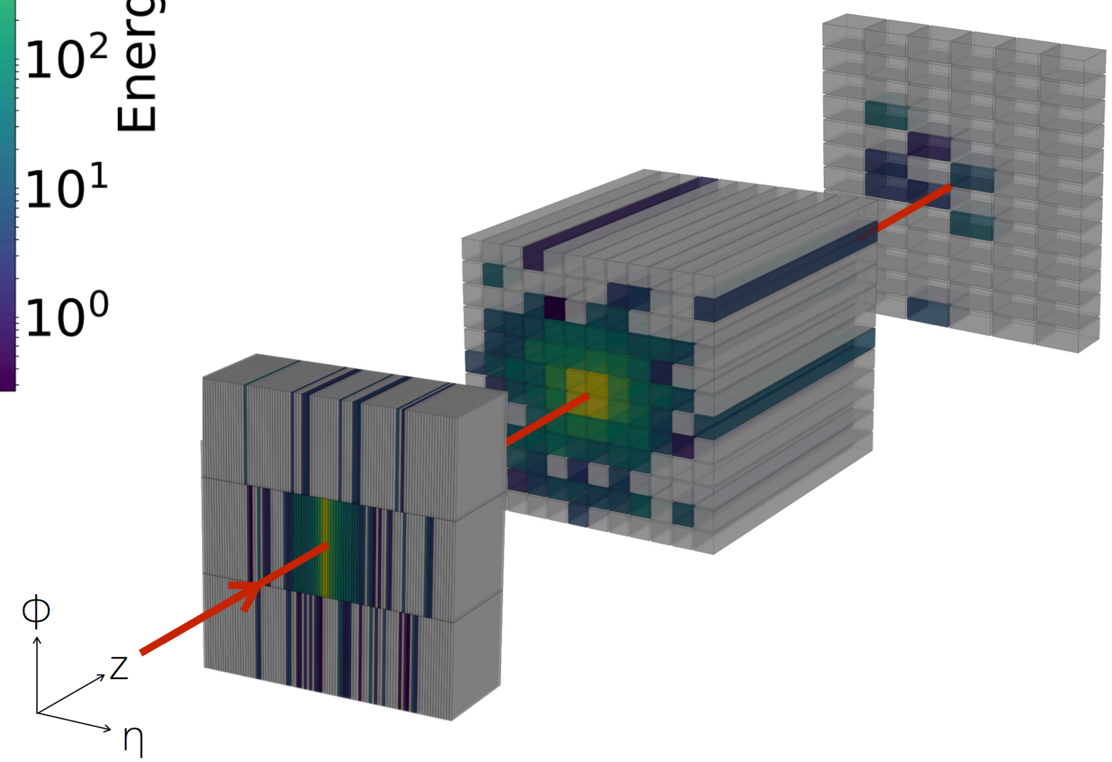
Image from 1705.02355



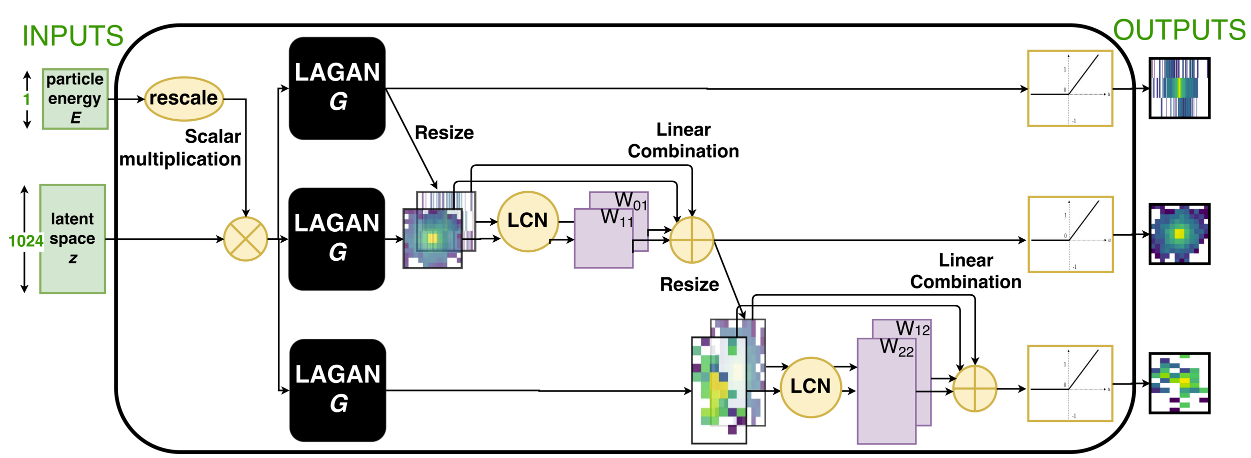
CNN methods became very popular for reconstruction/tagging tasks.

Detector simulation using ML

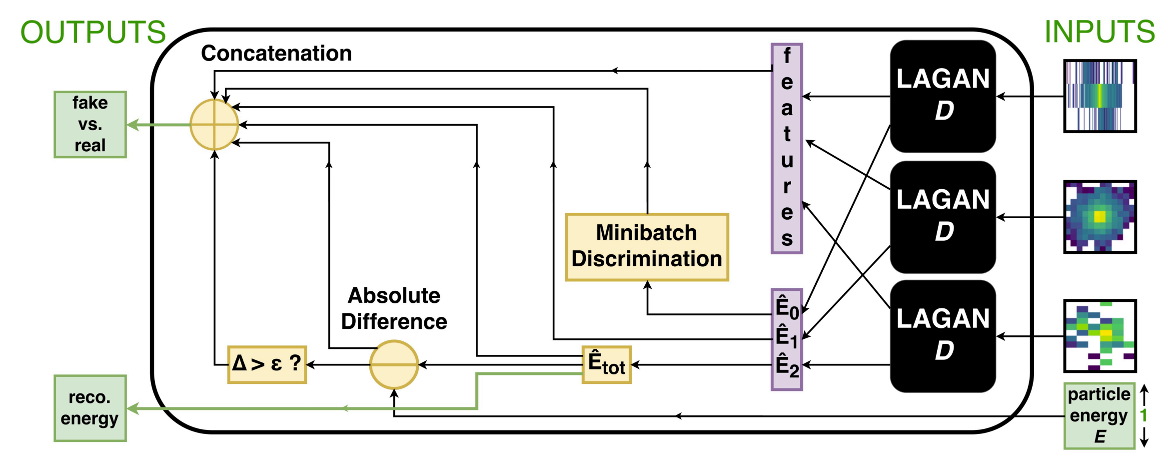




CaloGAN 1705.02355

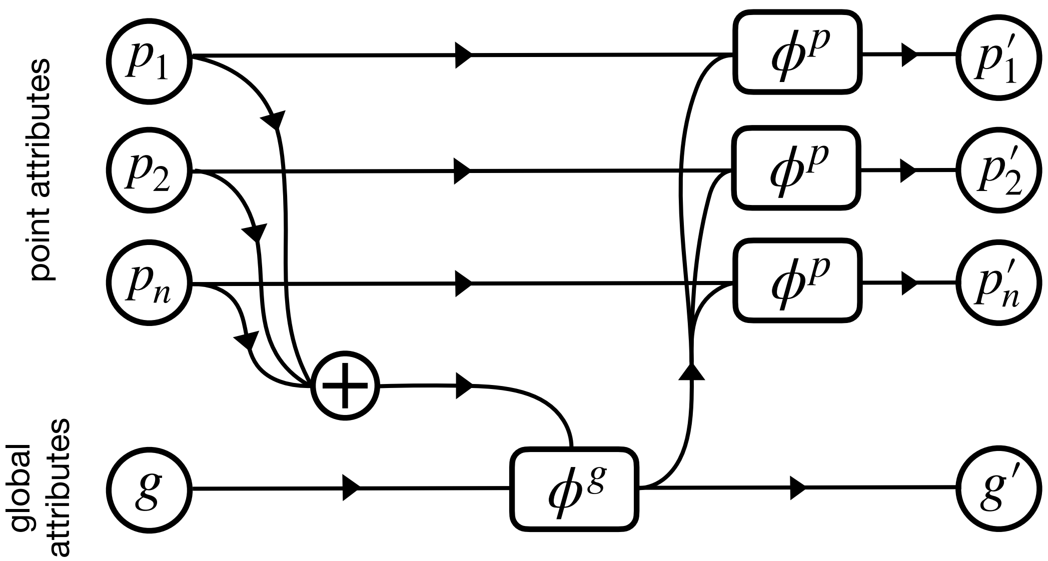


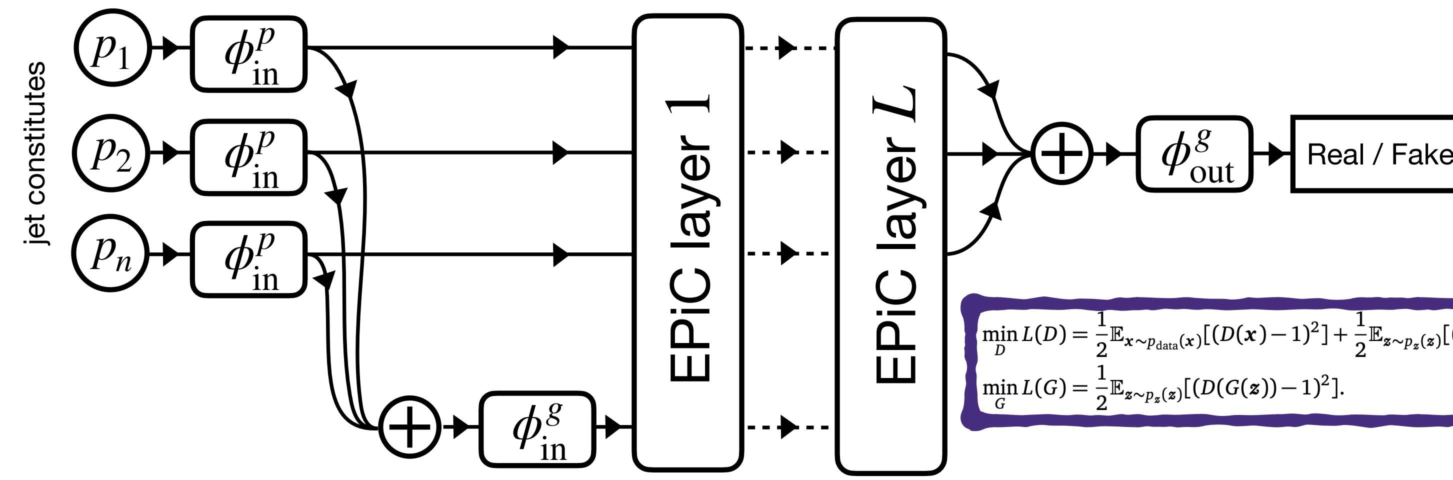
Generator



Discriminator

Detector simulation using ML

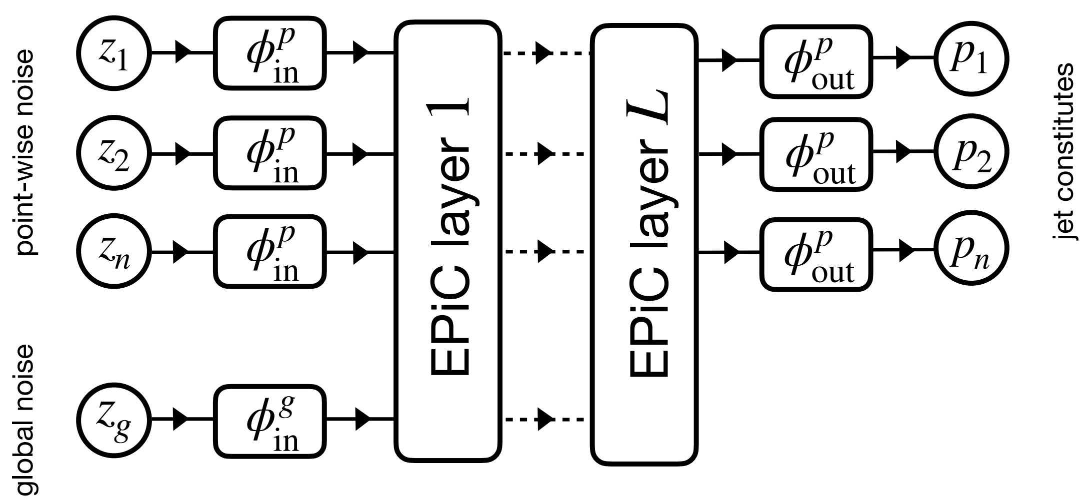


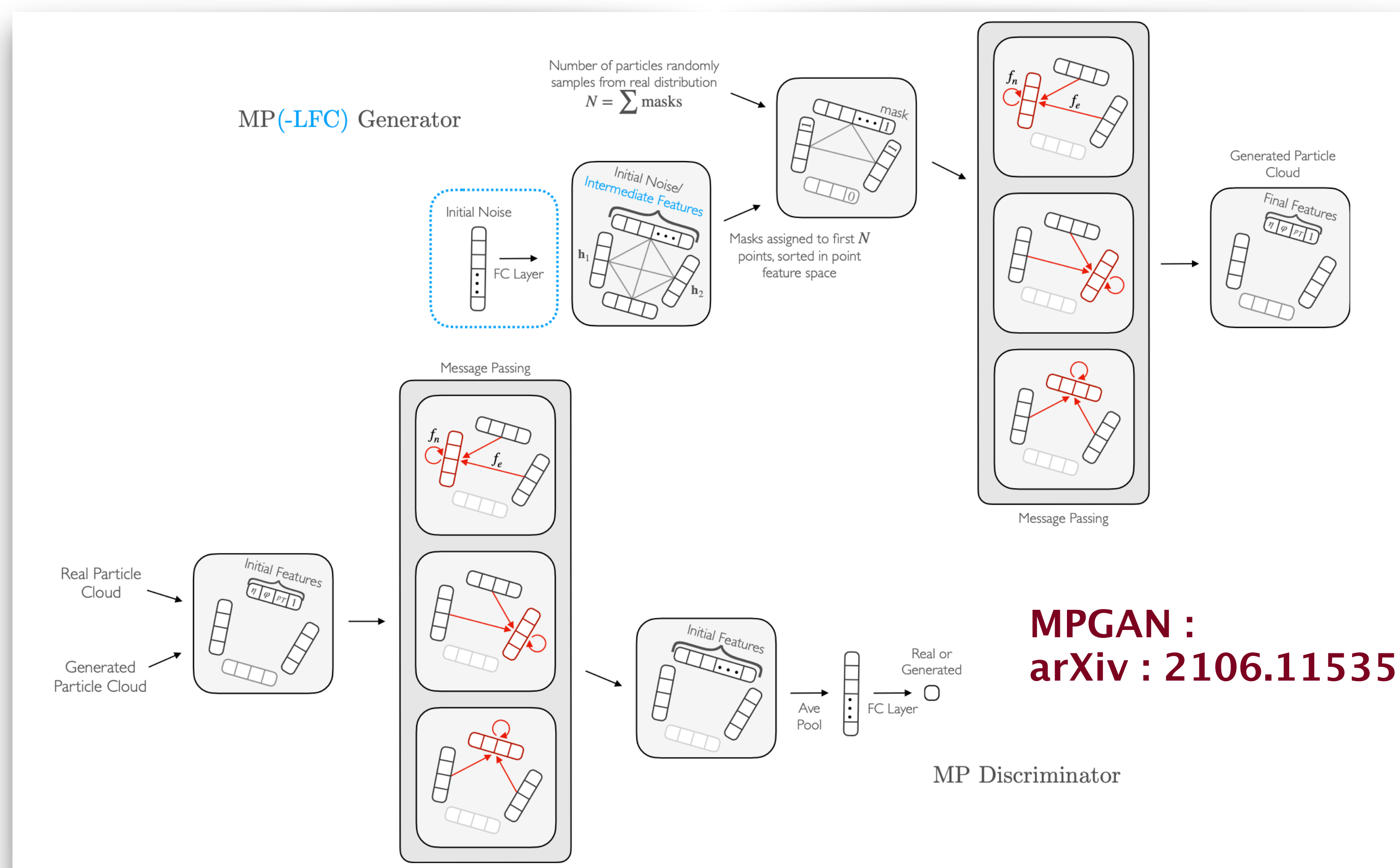


$$\min_D L(D) = \frac{1}{2}\mathbb{E}_{x\sim p_{\text{data}}(x)}[(D(x)-1)^2] + \frac{1}{2}\mathbb{E}_{z\sim p_z(z)}[(D(G(z)))^2],$$

$$\min_G L(G) = \frac{1}{2}\mathbb{E}_{z\sim p_z(z)}[(D(G(z))-1)^2].$$

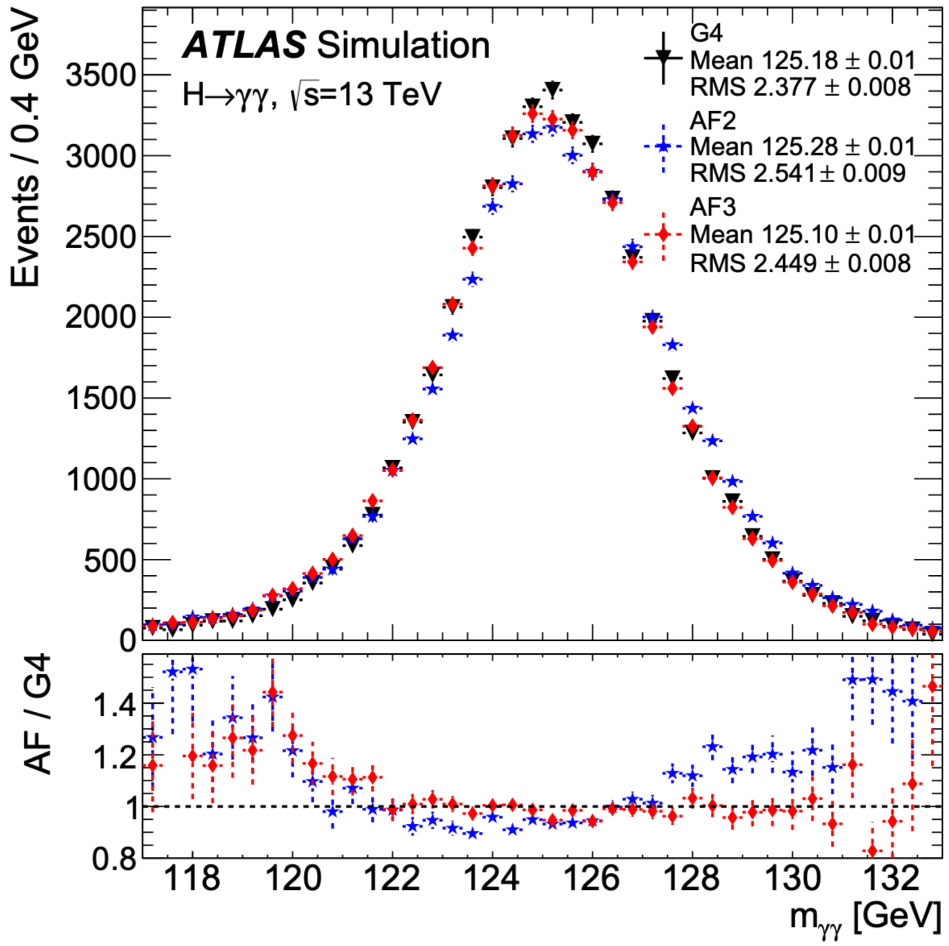
EPiC-GAN : SciPost Phys. 15, 130 (2023)

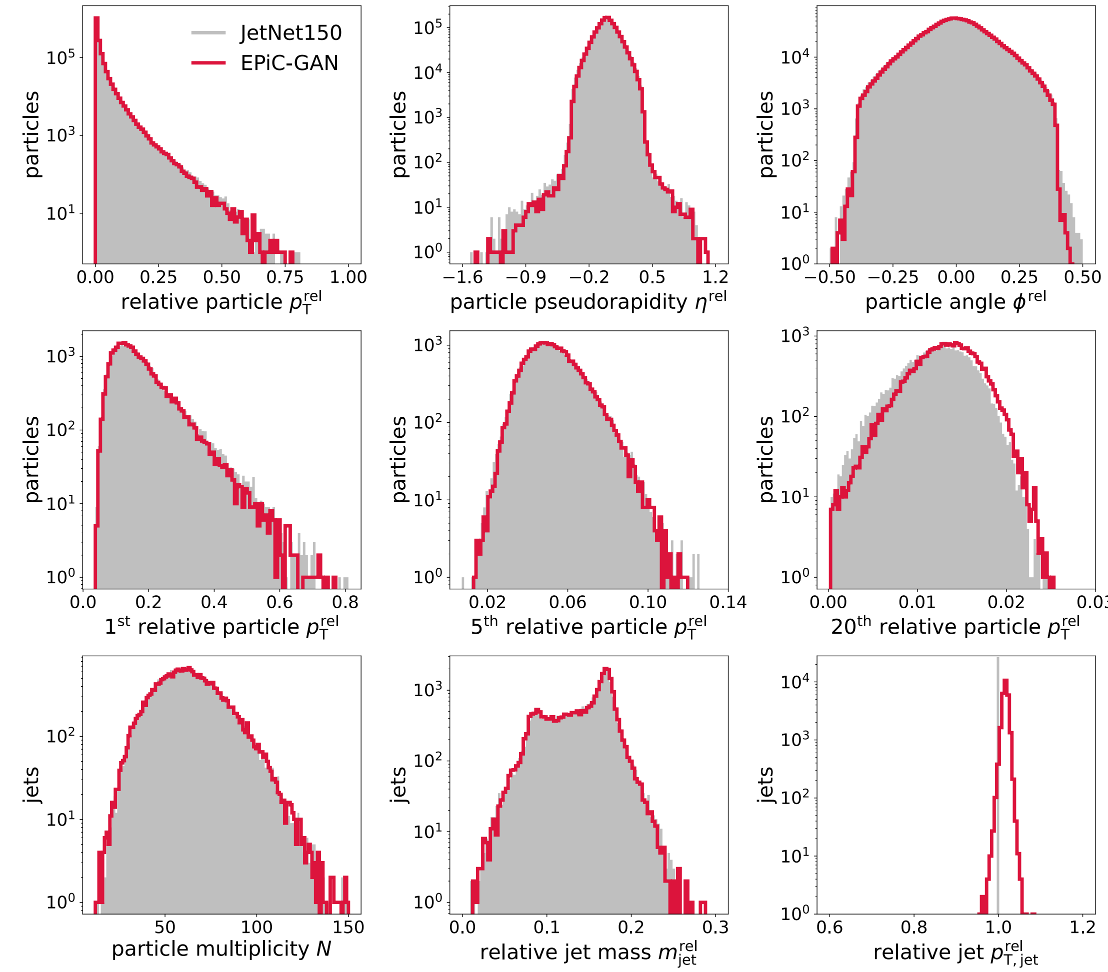




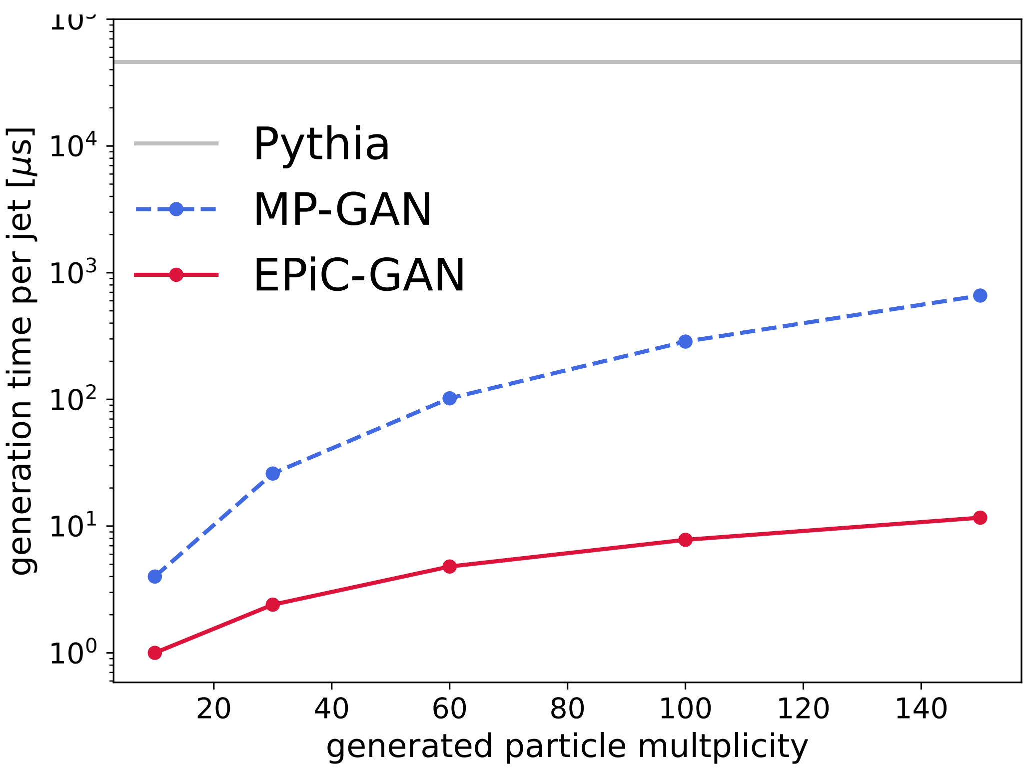
MPGAN : arXiv : 2106.11535

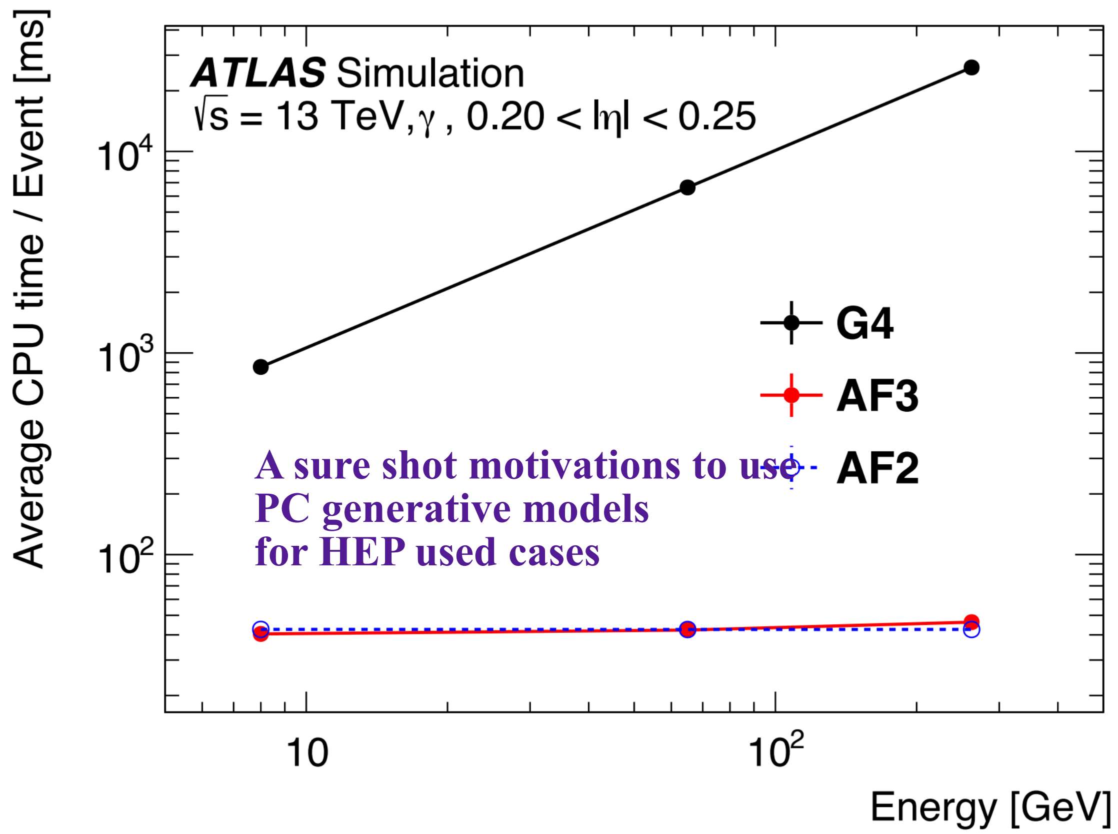
The major gain





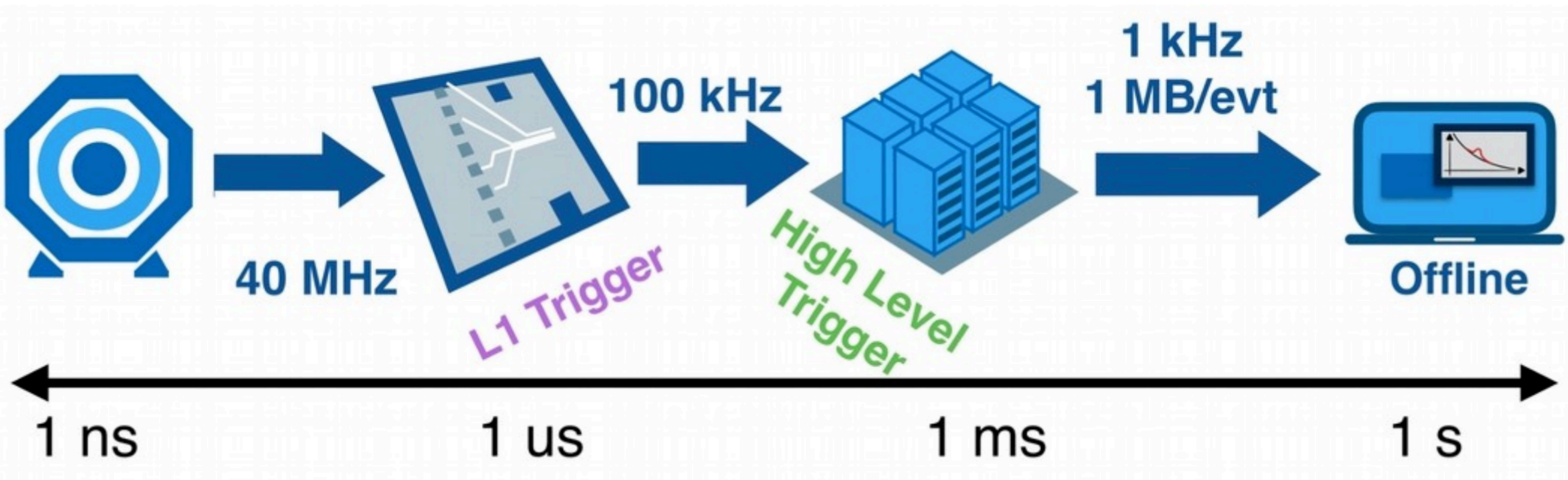
arXiv : 2109.02551





A sure shot motivations to use PC generative models for HEP used cases

Improving trigger using ML



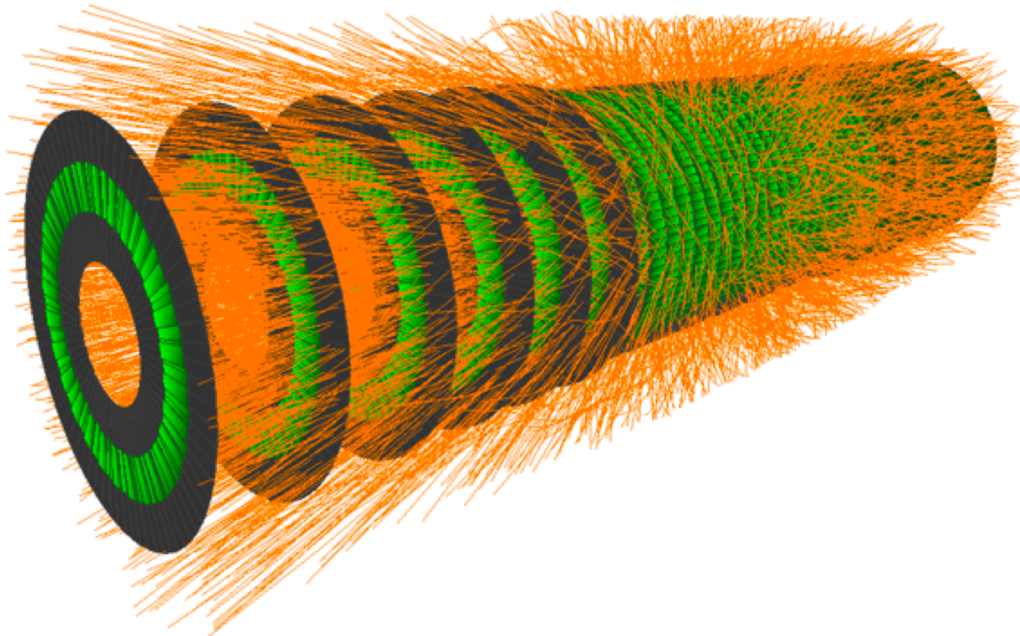
J. Duarte *et al* 2018 *JINST* 13 P07027

- L1 Trigger (hardware: FPGAs)
  - O(μs) hard latency. Typically coarse selection, BDT used for muon p_T assignment
- HLT (software: CPUs)
  - O(100 ms) soft latency. More complex algorithms (full detector information available), some BDTs and DNNs used
- Offline (software: CPUs)
  - > 1 s latencies. Full event reconstruction, bulk of machine learning usage in CMS

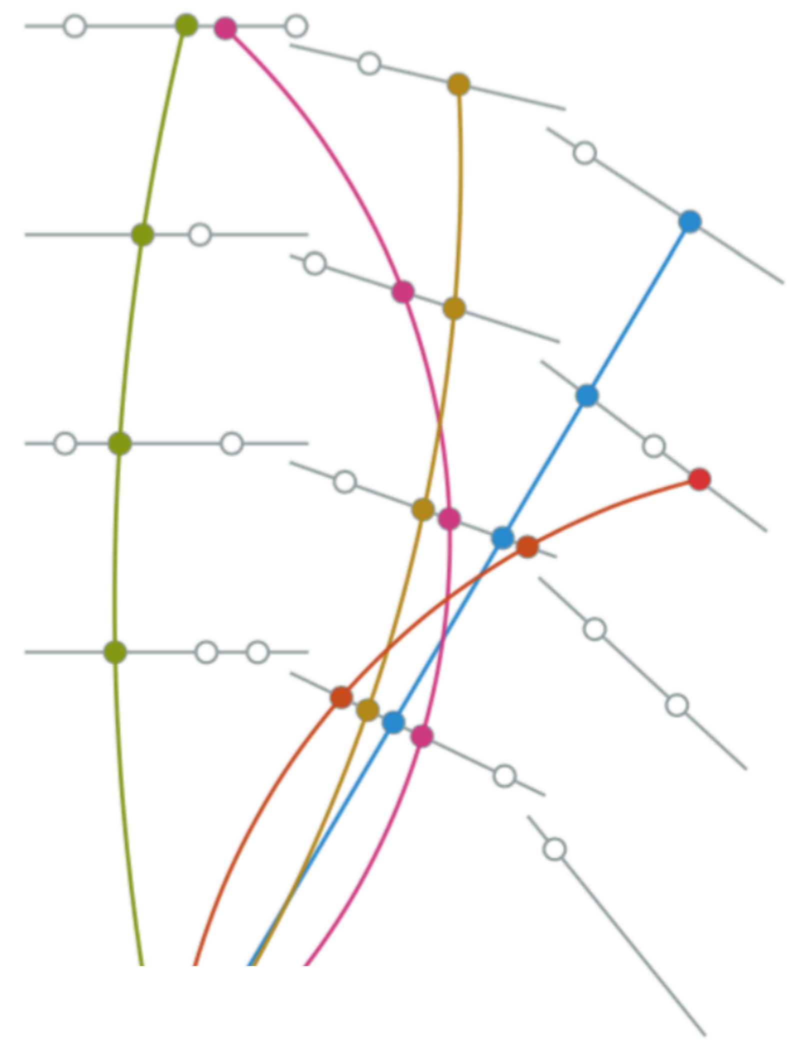
hls 4 ml

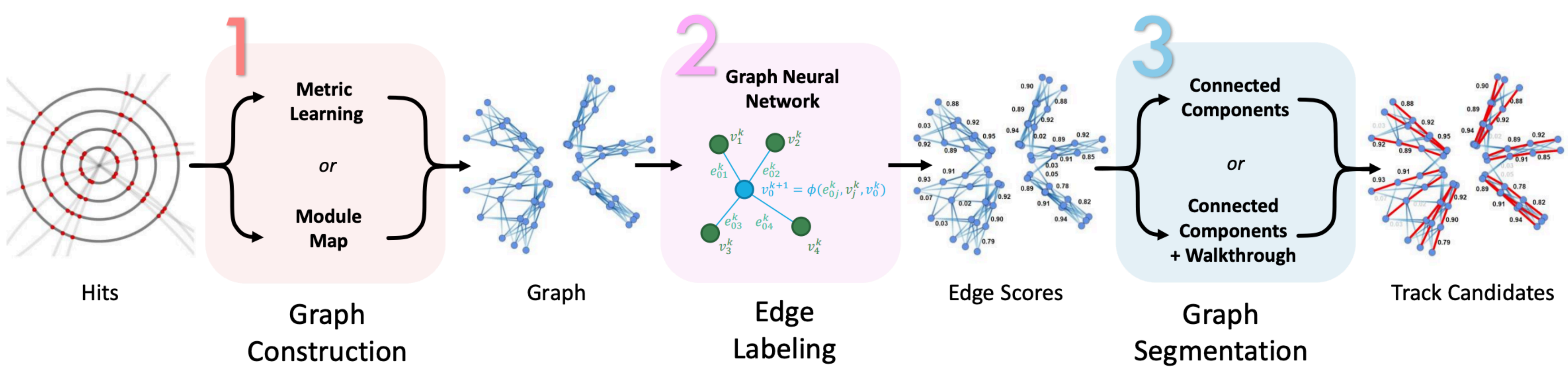
Original slide

Tracking & ML

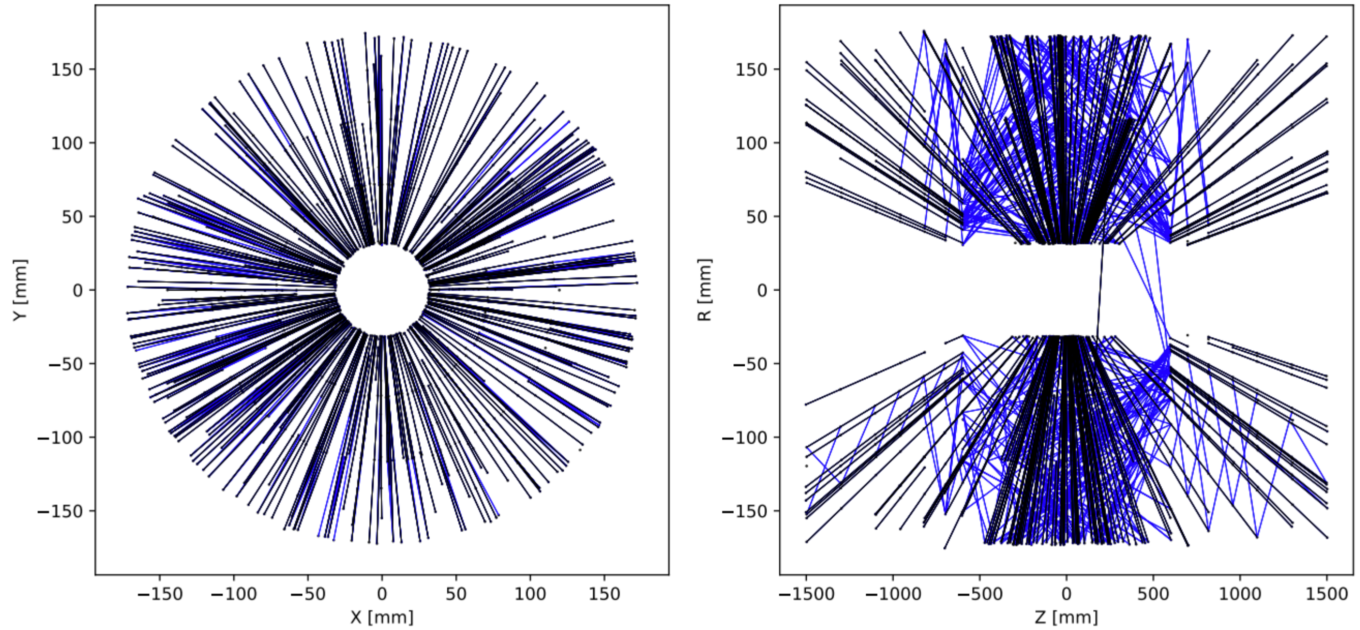


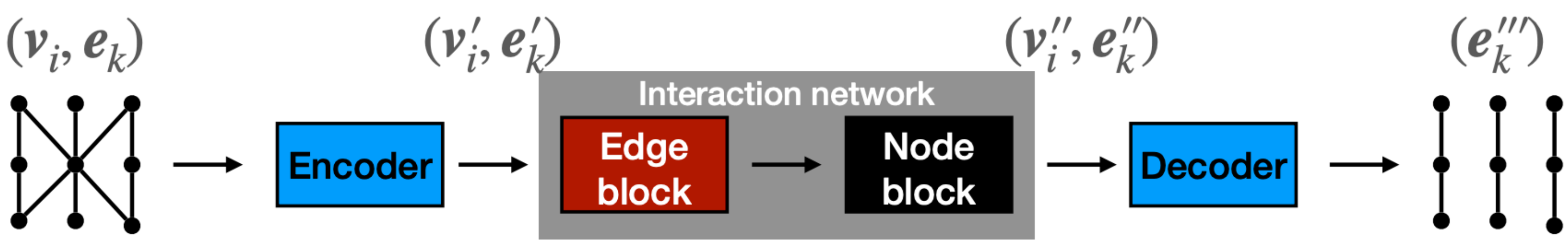
An exponentially large edge finding problem





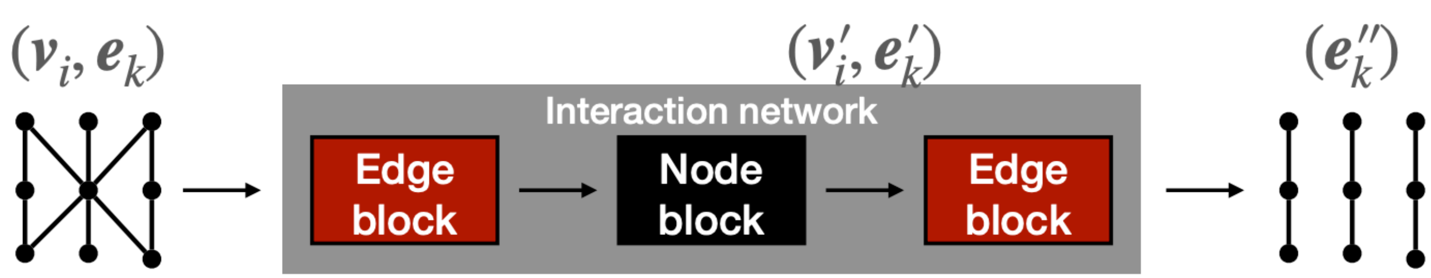
Tracking & ML





$e_k' = \phi_1^e(e_k)$ $e_k'' = \phi_2^e(e_k', v_{r_k}', v_{s_k}')$ $v_i'' = \phi_2^v(\bar{e}_i'', v_i')$ $e_k''' = \phi_3^e(e_k'')$

$v_i' = \phi_1^v(v_i)$ $\bar{e}_i'' = \rho^{e\to v}(E_i')$

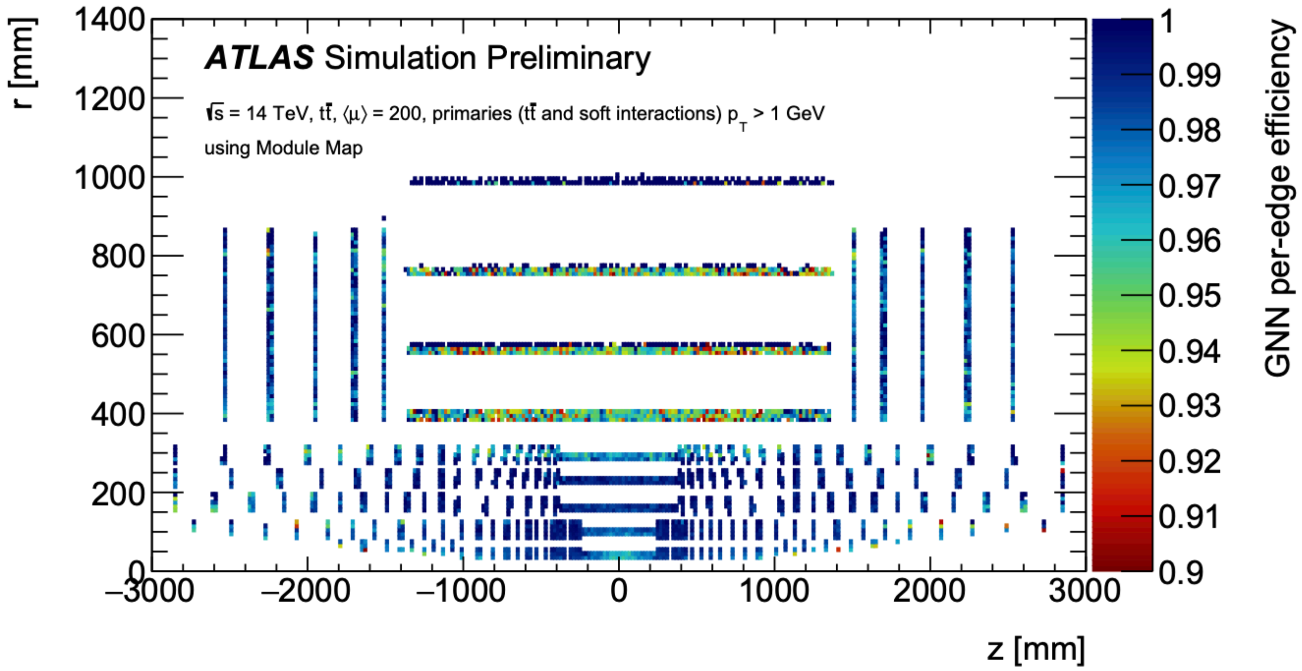


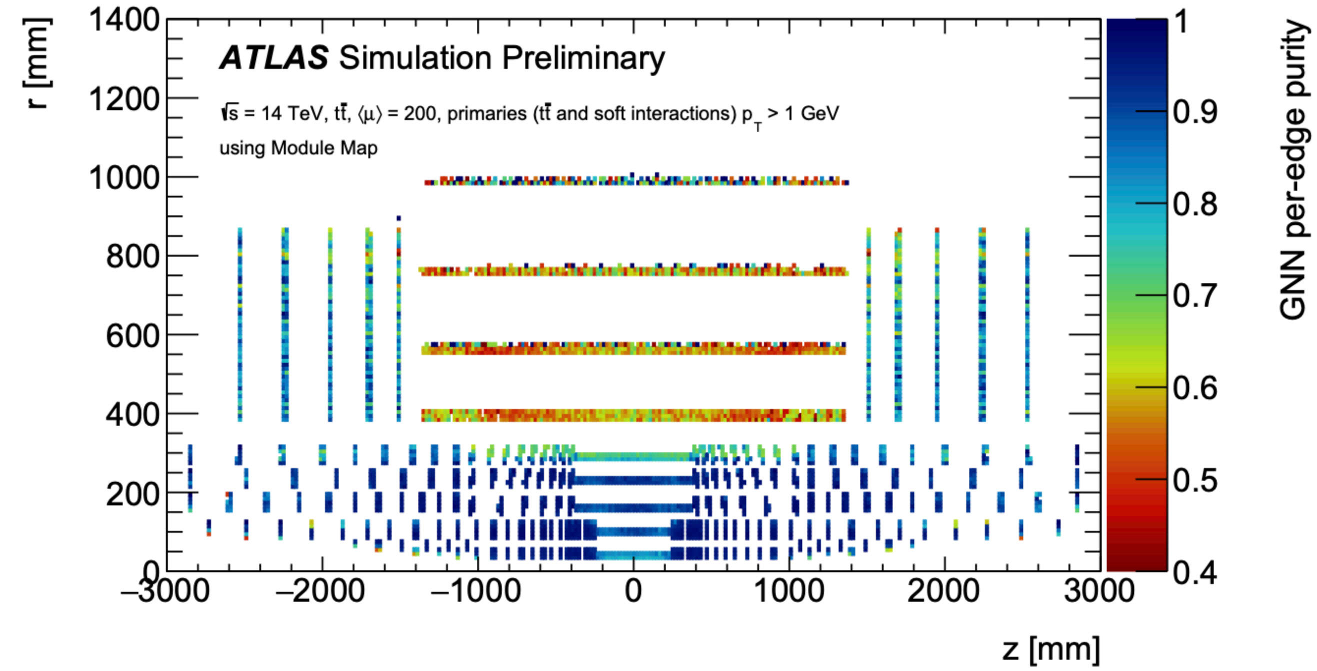
$e_k' = \phi_2^e(e_k, v_{r_k}, v_{s_k})$ $v_i' = \phi_2^v(\bar{e}_i', v_i)$ $e_k'' = \phi_2^e(e_k', v_{r_k}', v_{s_k}')$

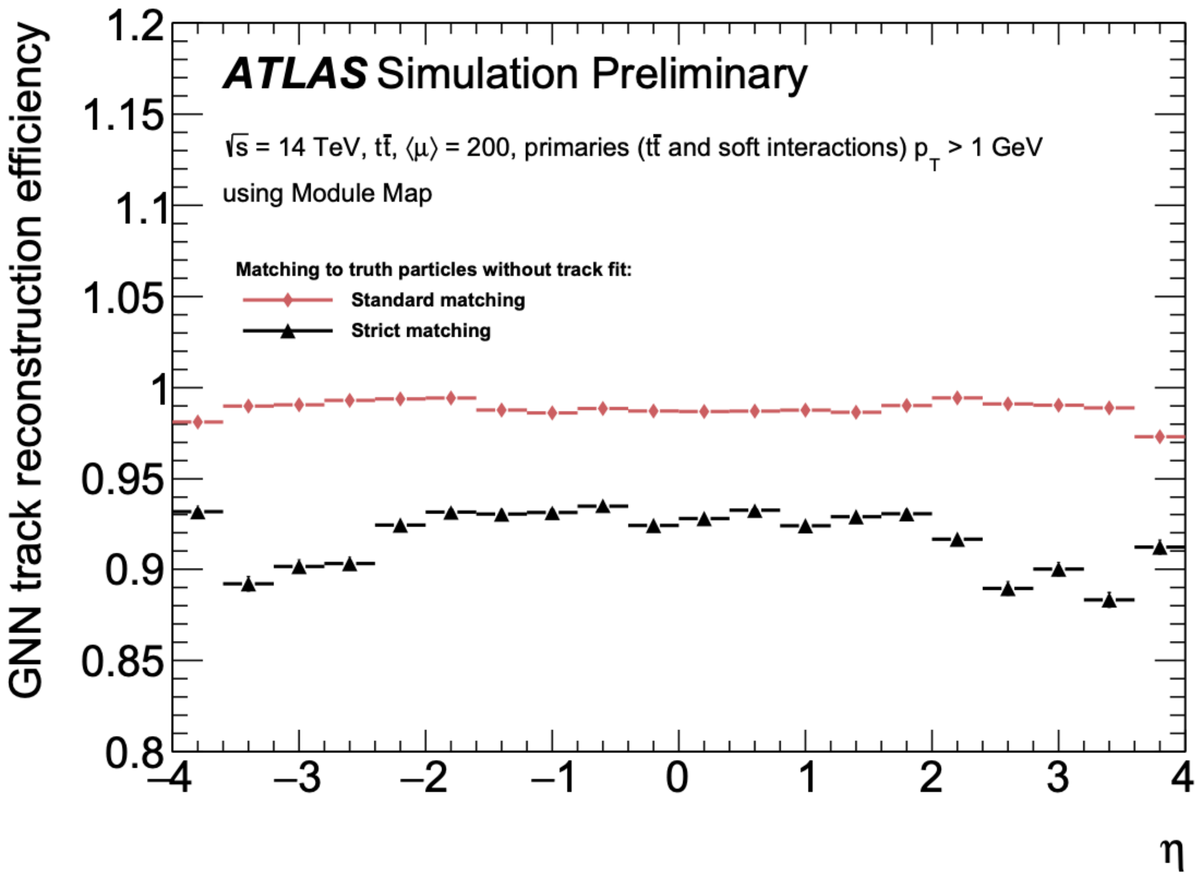
$\bar{e}_i' = \rho^{e\to v}(E_i)$

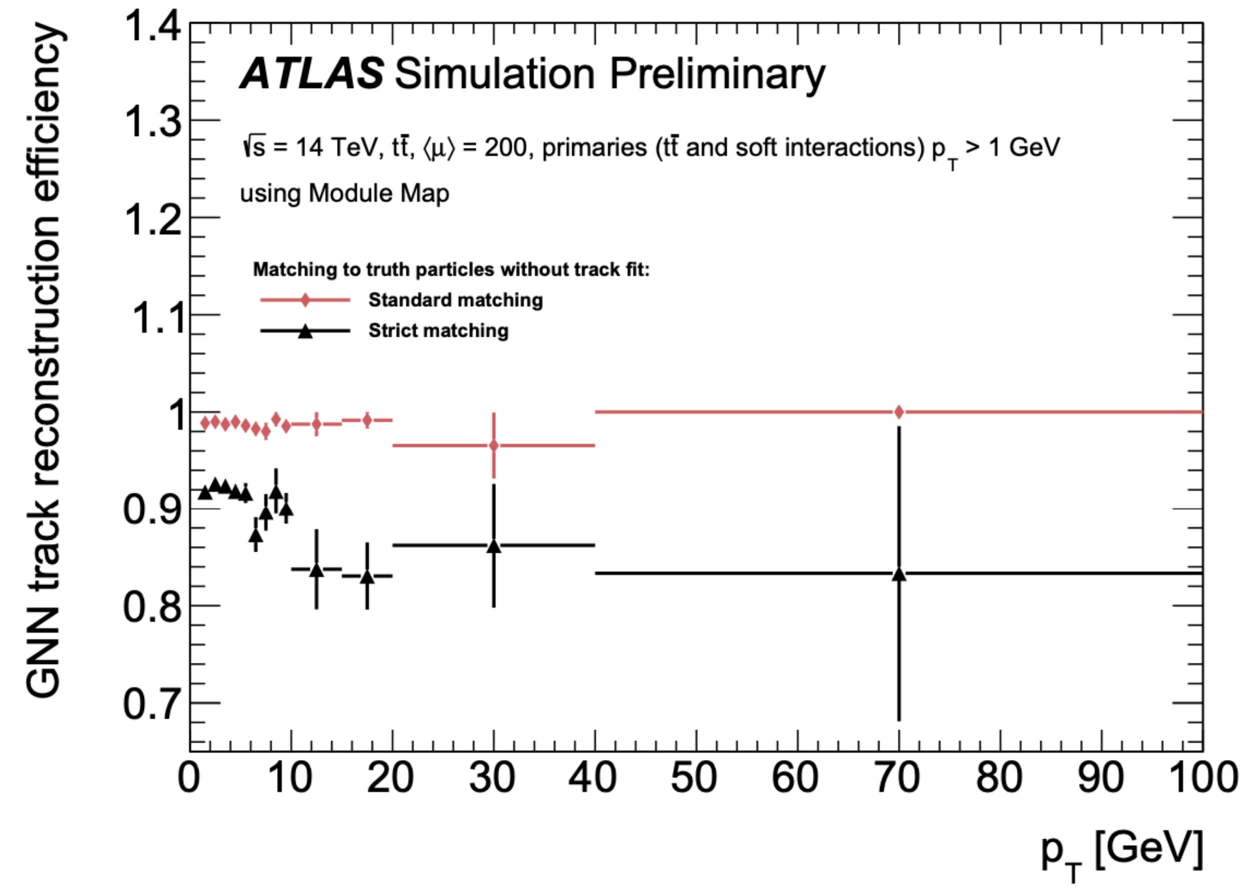
Tracking & ML

ATL-ITK-PROC-2022-006

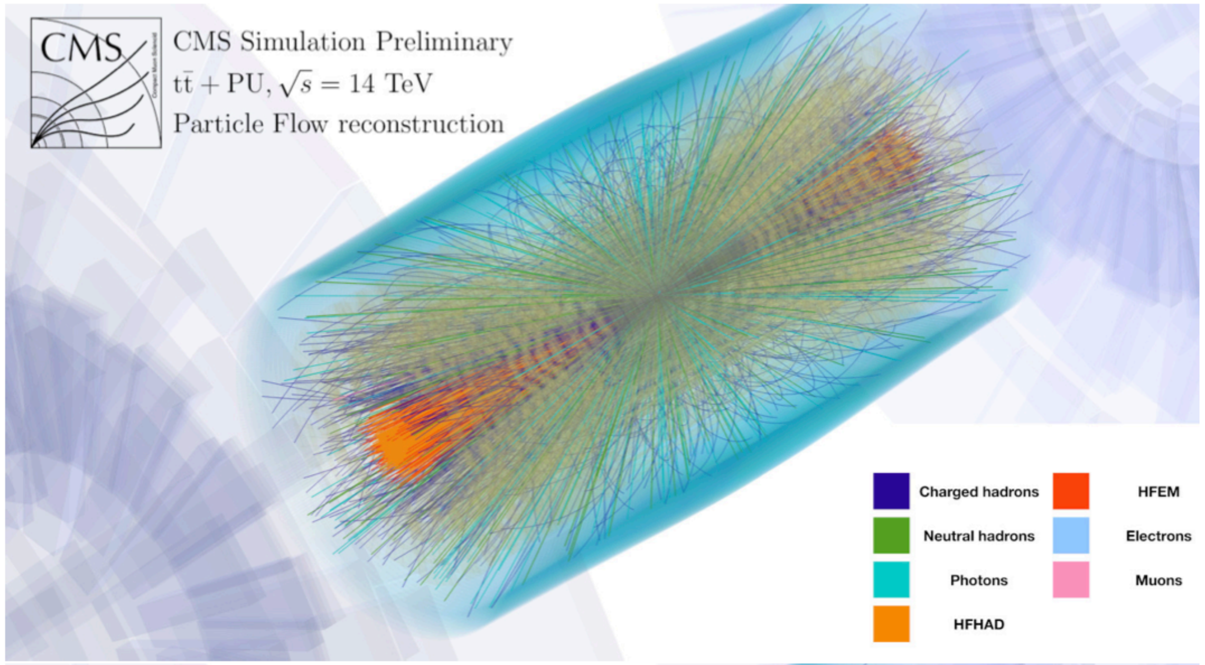


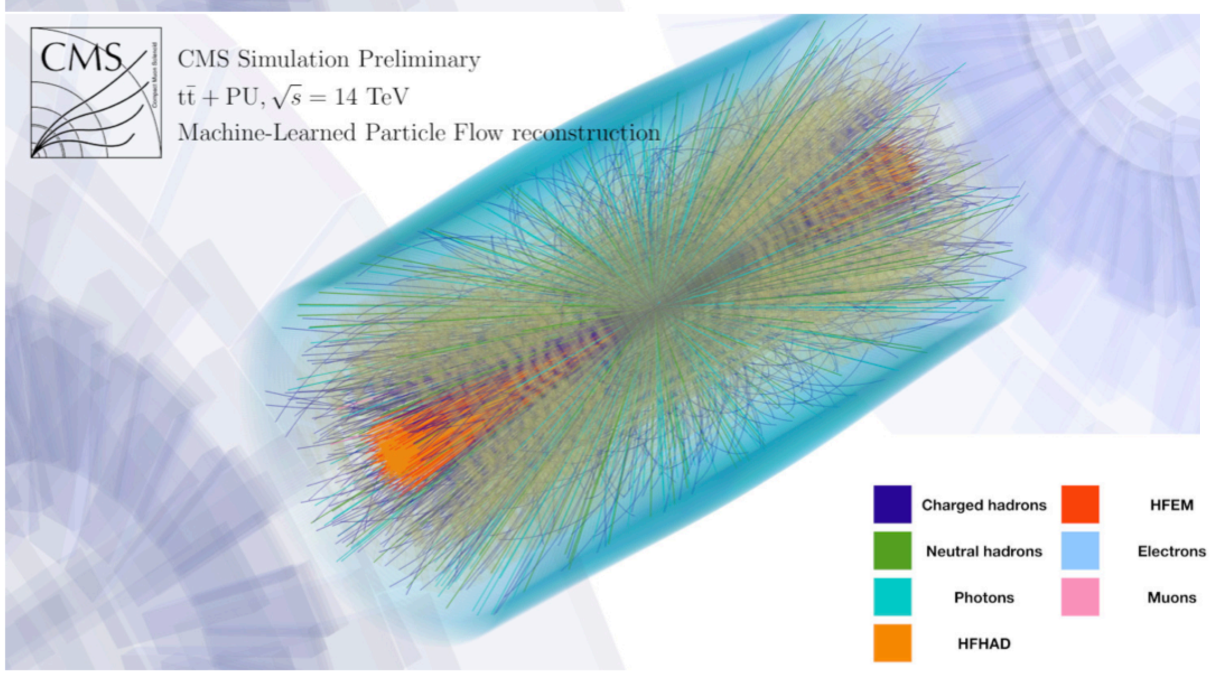


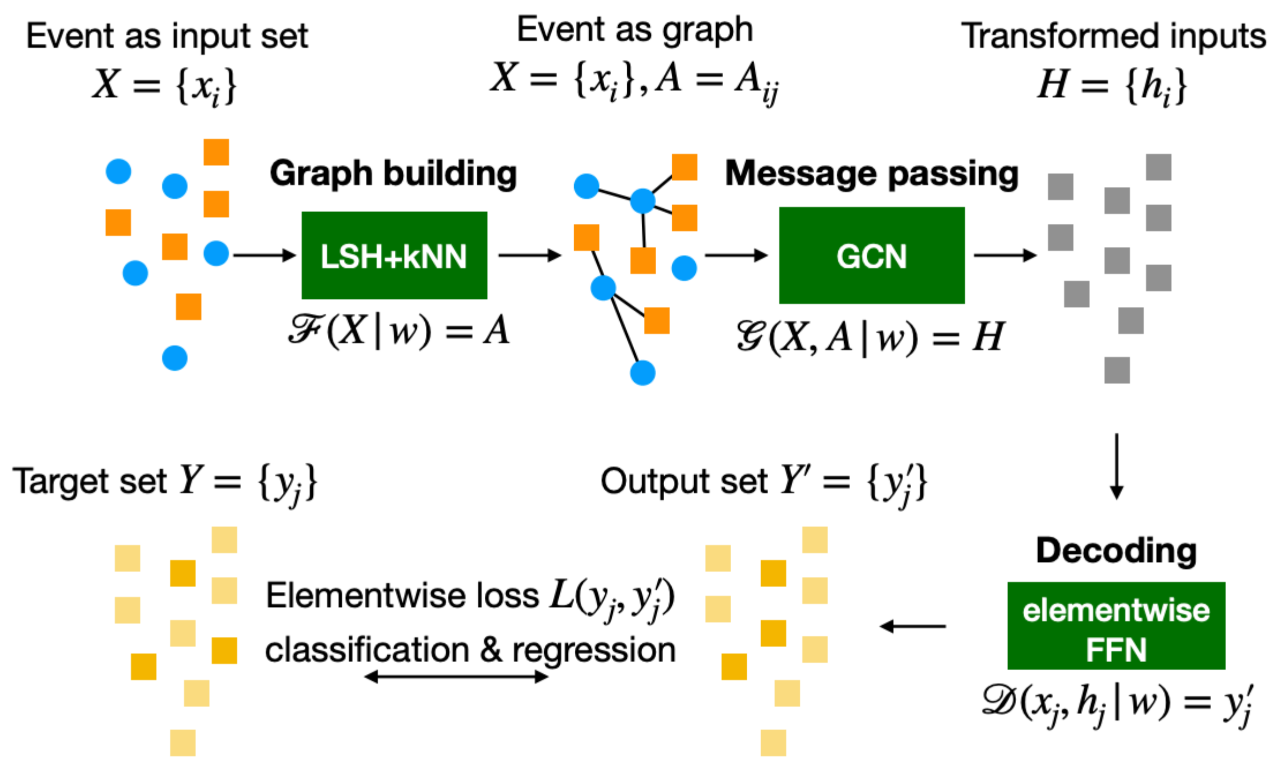




Full ML driven PFlow : MLPF







$x_i = [\text{type}, p_T, E_{ECAL}, E_{HCAL}, \eta, \phi, \eta_{outer}, \phi_{outer}, q, \ldots], \quad \text{type} \in \{\text{track}, \text{cluster}\}$

$y_j = [\text{PID}, p_T, E, \eta, \phi, q, \ldots], \quad \text{PID} \in \{\text{none}, \text{charged hadron}, \text{neutral hadron}, \gamma, e^\pm, \mu^\pm\}$

$h_i \in \mathbb{R}^{256}$

Trainable neural networks: $\mathscr{F}, \mathscr{G}, \mathscr{D}$

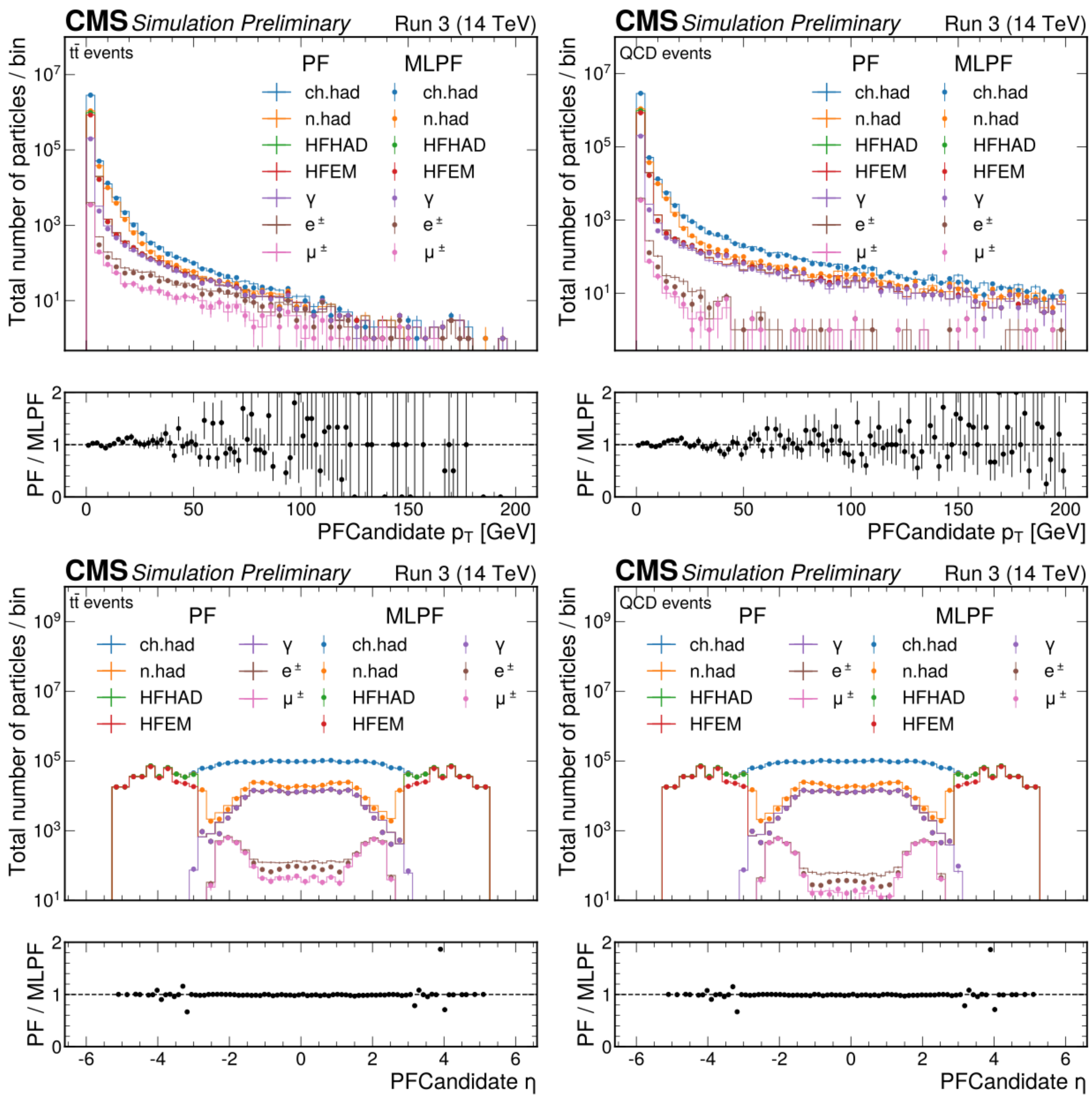
● - track, ■ - calorimeter cluster, ■ - encoded element

■ - target (predicted) particle, ■ - no target (predicted) particle

MLPF
Eur. Phys. J. C (2021) 81: 381
J. Pata et. al.

PF lepton, hadron, photon = F_{PF} (track hits + calo cells)

Combining track + calo for PFlow



MLPF

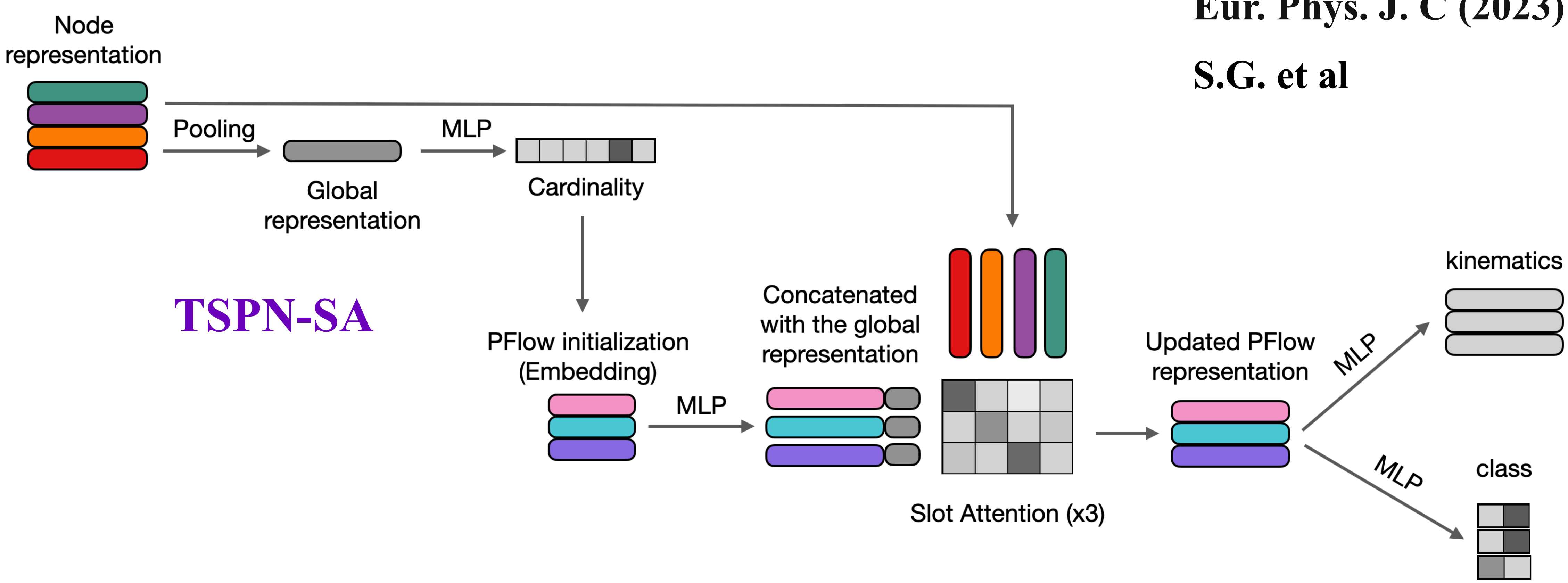
J. Phys.: Conf. Ser. 2438, 012100 (2023)

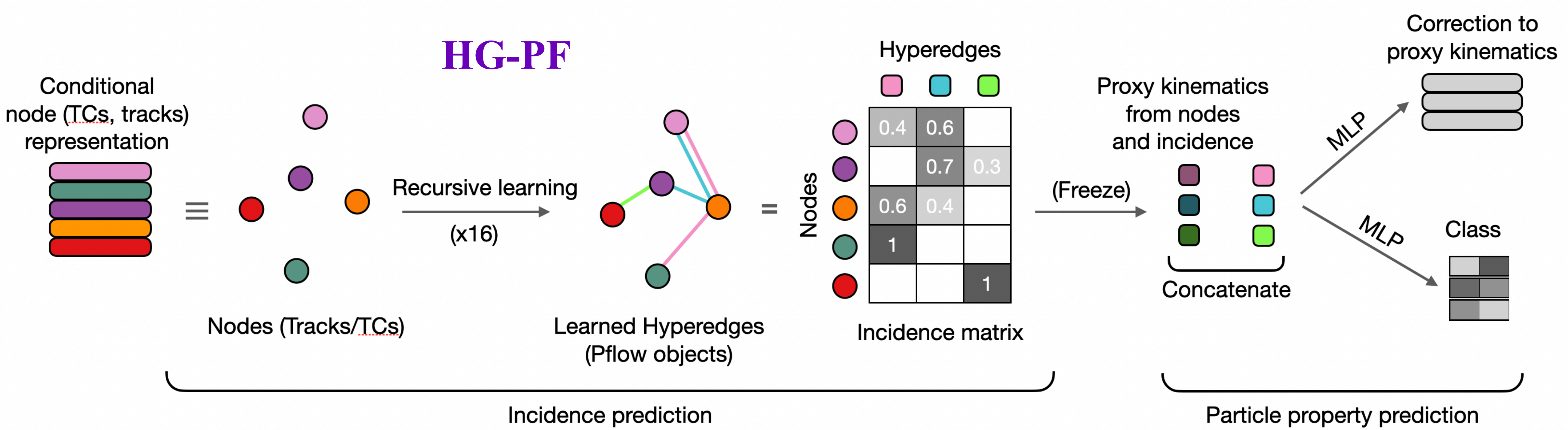
J. Pata et. al.

Attempt for higher order correlation

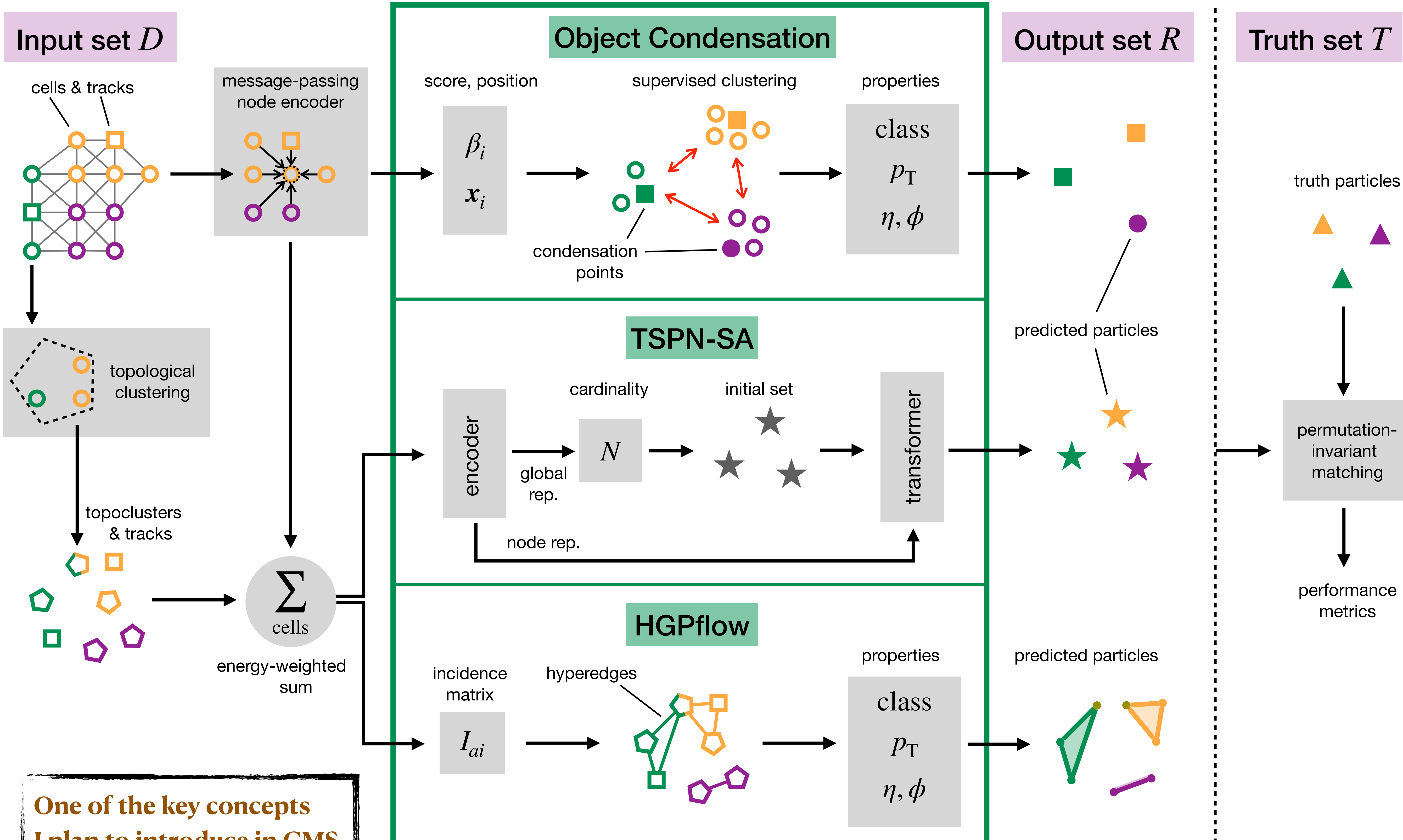
Eur. Phys. J. C (2023) 83:596

S.G. et al



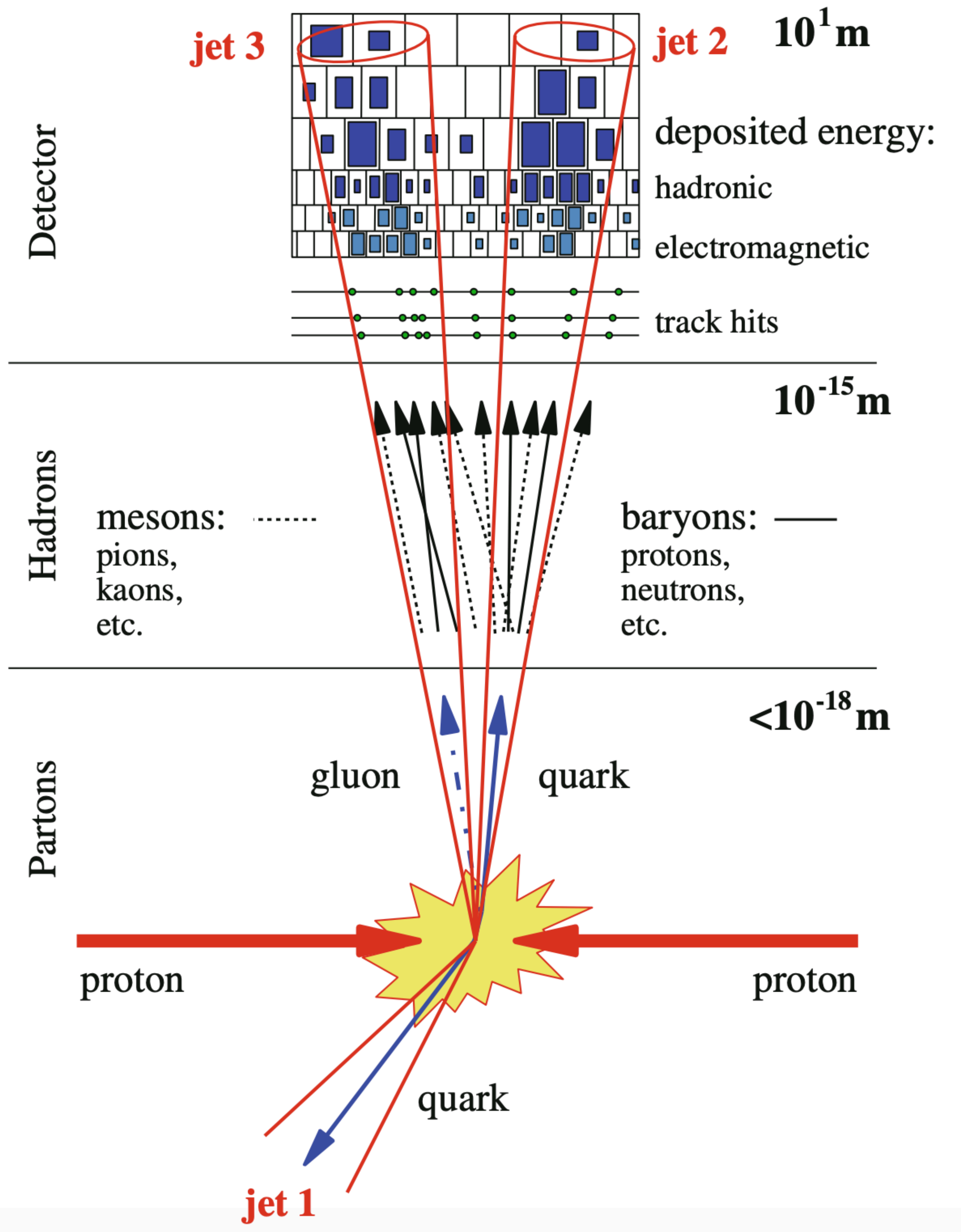


The network flow comparisons



One of the key concepts
I plan to introduce in CMS

What do we want to tag : Jets 101



The Large Hadron Collider
Editors: Thomas Schörner-Sadenius

$$\{p_1, p_2, \ldots, p_n\}$$

$$\downarrow \text{ Jet Algorithm (for CA, kT, anti-kT)}$$

$$\{j_1, j_2, \ldots, j_k\}$$

$$\{p_1, p_2, \ldots, p_n\} = F(q)$$

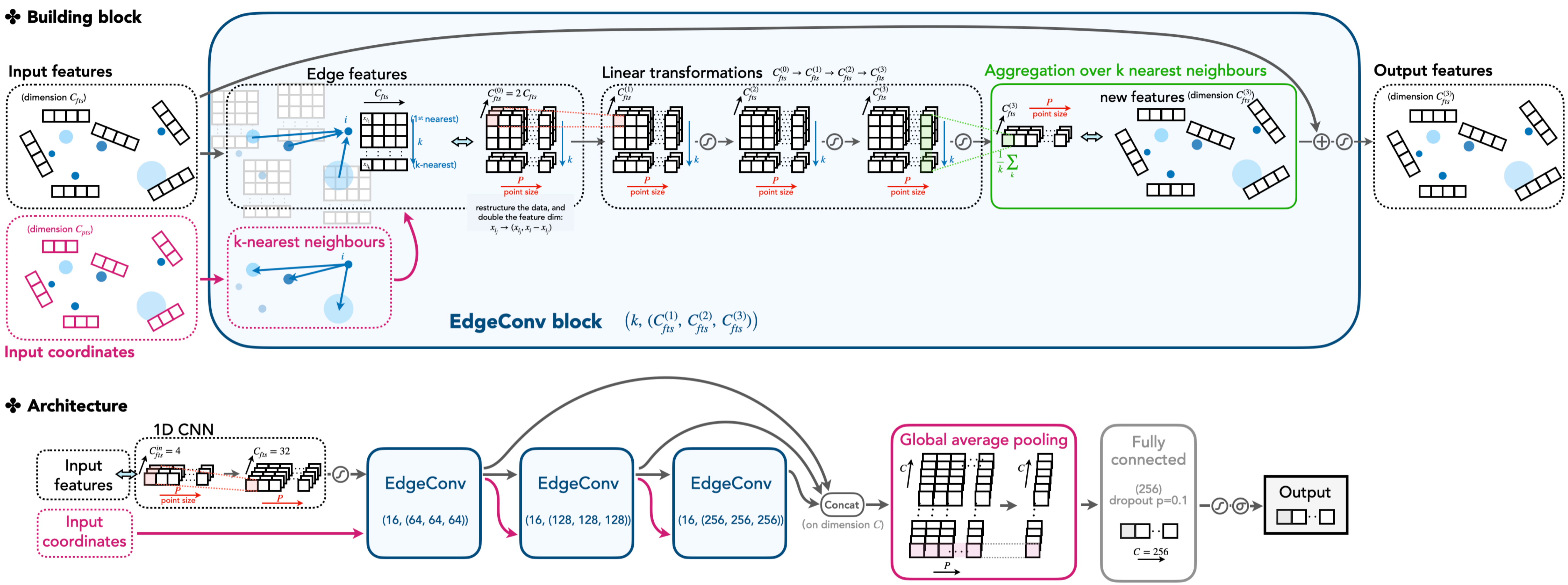
The forward problem is not computable from first principle

The question of jet tagging is how do we define the inverse problem?

$$q = F^{-1}\Big(\{p_1, p_2, \ldots, p_n\}\Big) \; ?$$

Object tagging

Particle Net : 1902.08570



arXiv > cs > arXiv:1801.07829

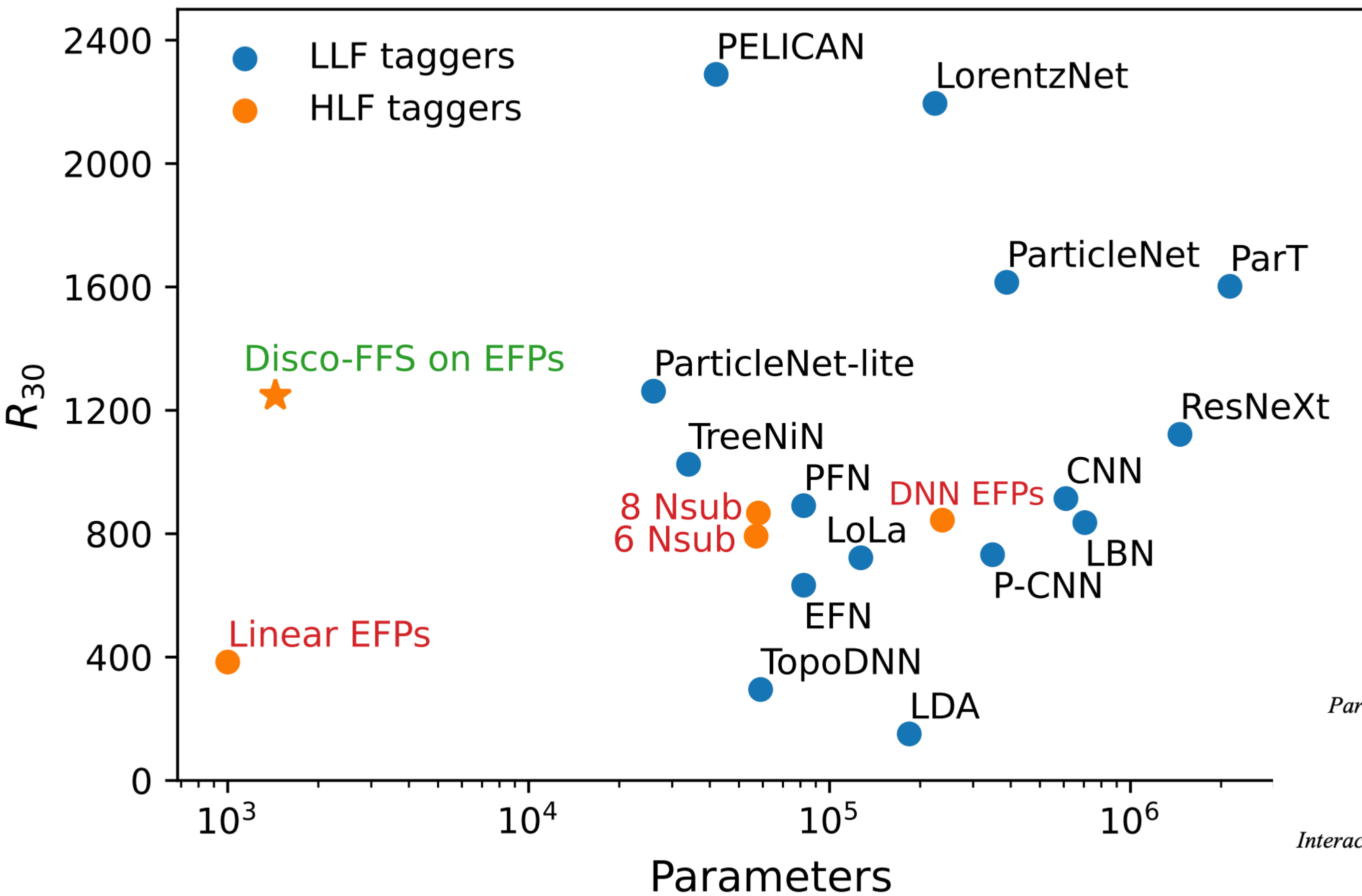
Computer Science > Computer Vision and Pattern Recognition

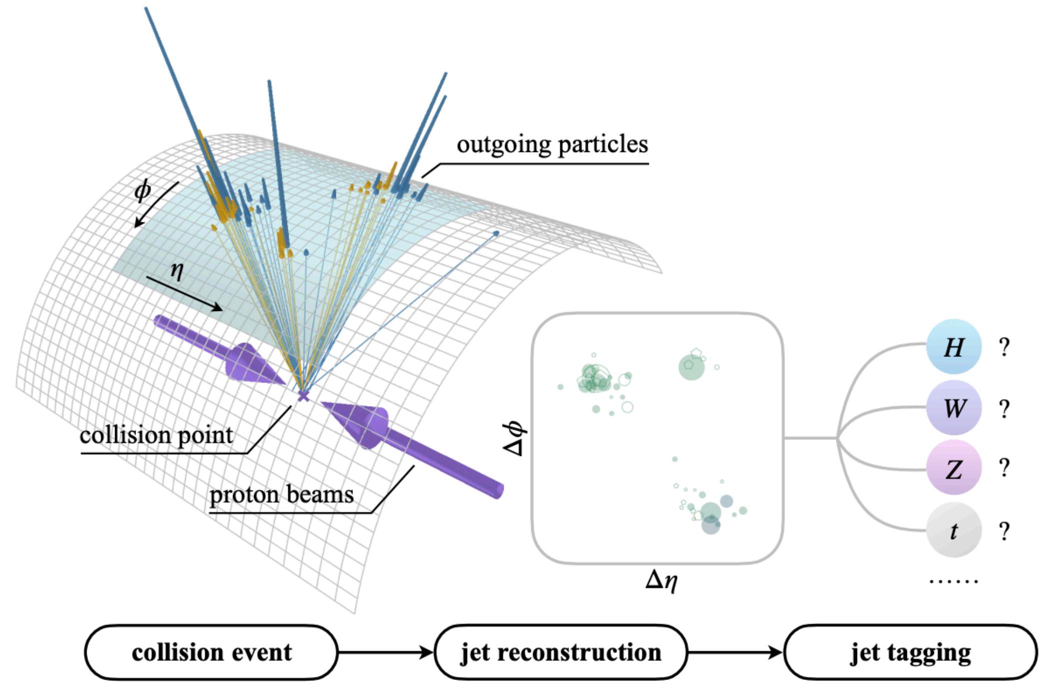
[Submitted on 24 Jan 2018 (v1), last revised 11 Jun 2019 (this version, v2)]

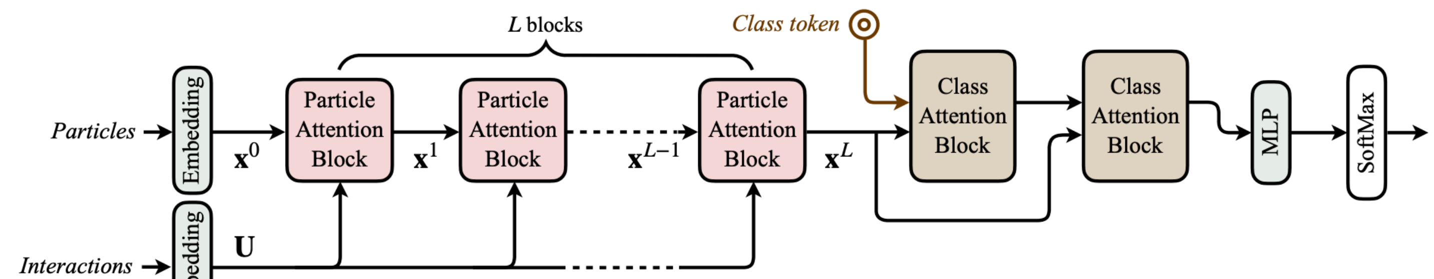
Dynamic Graph CNN for Learning on Point Clouds

Yue Wang, Yongbin Sun, Ziwei Liu, Sanjay E. Sarma, Michael M. Bronstein, Justin M. Solomon

Object tagging



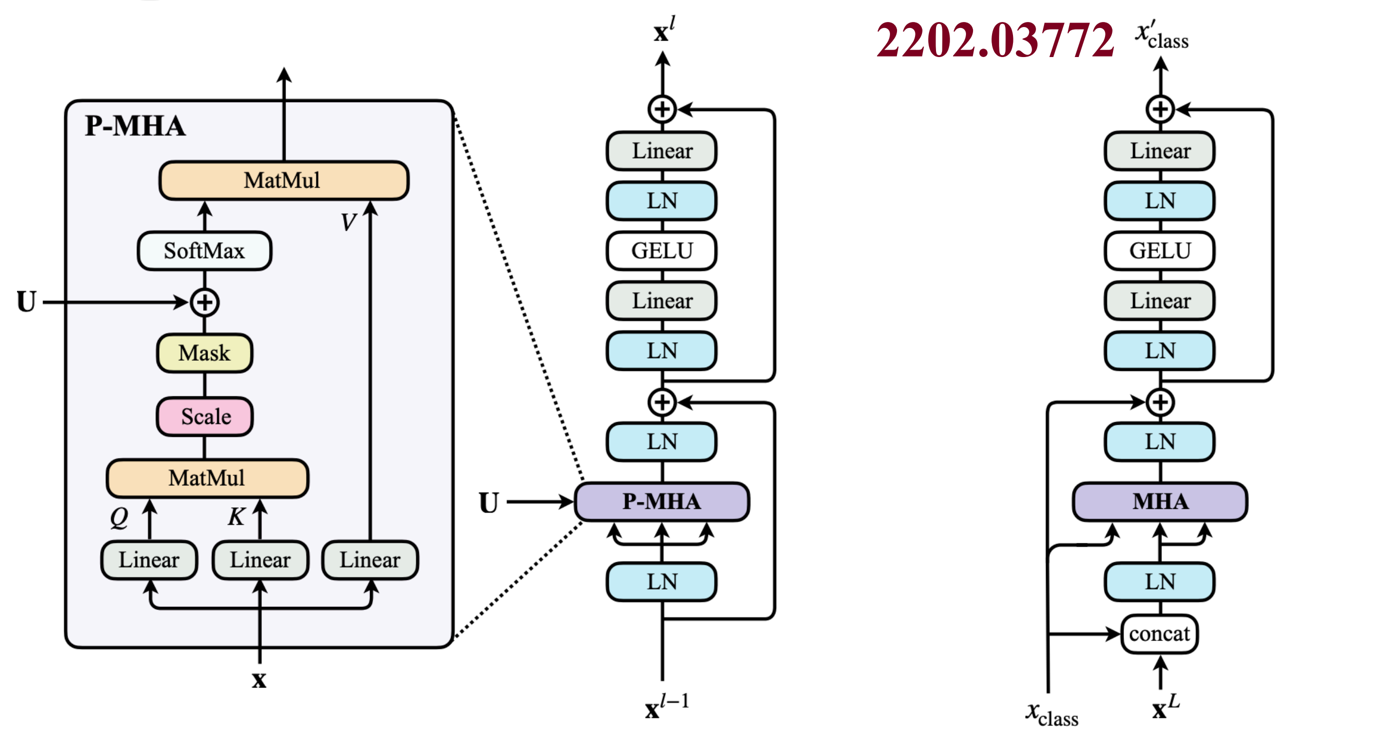




(a) Particle Transformer

Particle Transformer : 2202.03772

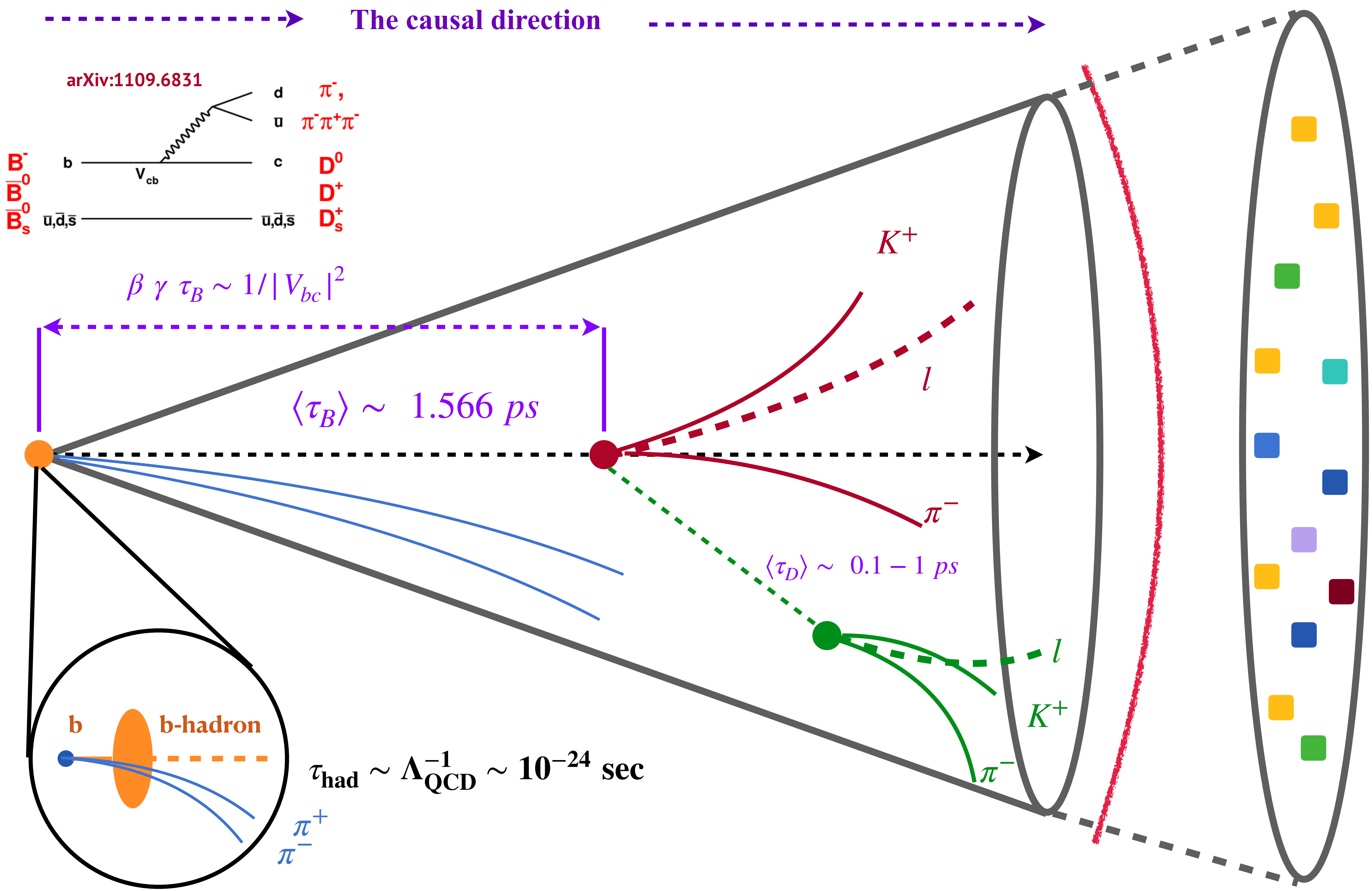
	Accuracy	AUC	$\text{Rej}_{50\%}$	$\text{Rej}_{30\%}$
P-CNN	0.930	0.9803	201 ± 4	759 ± 24
PFN	—	0.9819	247 ± 3	888 ± 17
ParticleNet	0.940	0.9858	397 ± 7	1615 ± 93
JEDI-net (w/ $\sum O$)	0.930	0.9807	—	774.6
PCT	0.940	0.9855	392 ± 7	1533 ± 101
LGN	0.929	0.964	—	435 ± 95
rPCN	—	0.9845	364 ± 9	1642 ± 93
LorentzNet	0.942	0.9868	498 ± 18	2195 ± 173
ParT	0.940	0.9858	413 ± 16	1602 ± 81
ParticleNet-f.t.	0.942	0.9866	487 ± 9	1771 ± 80
ParT-f.t.	**0.944**	**0.9877**	$\mathbf{691 \pm 15}$	$\mathbf{2766 \pm 130}$



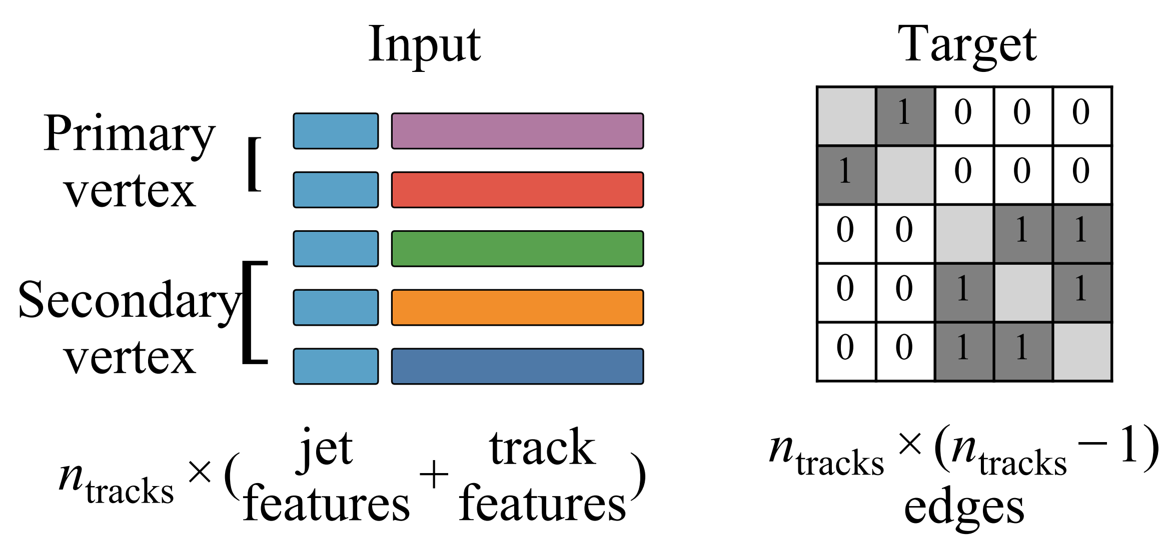
(b) Particle Attention Block

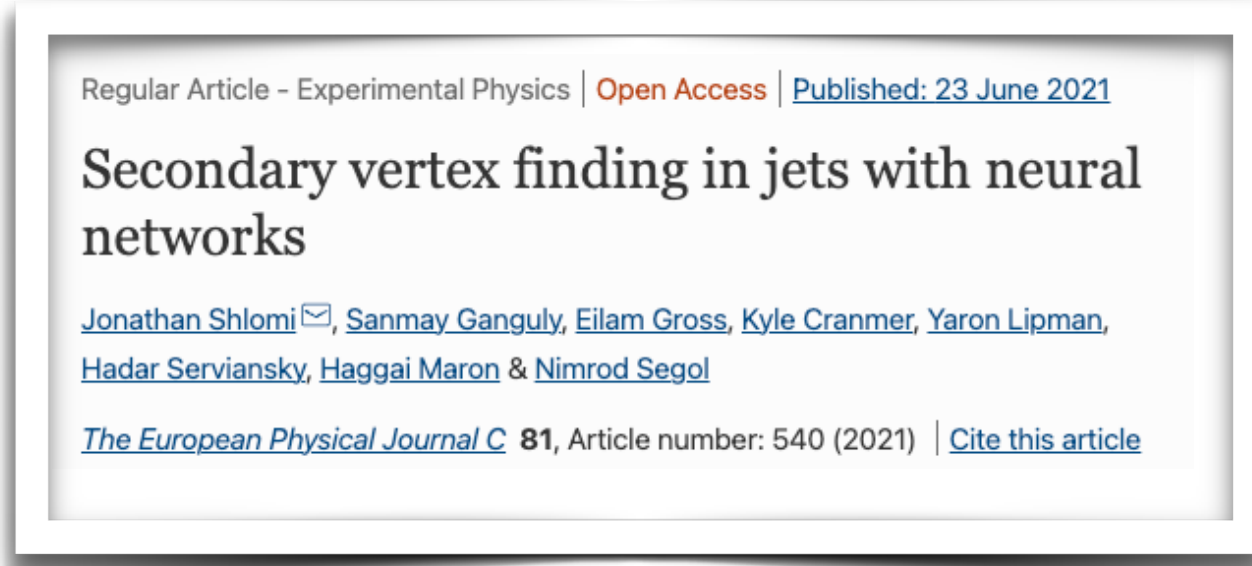
(c) Class Attention Block

Anatomy of heavy-quark hadronization



Set2Graph proposal for flavor-tagging

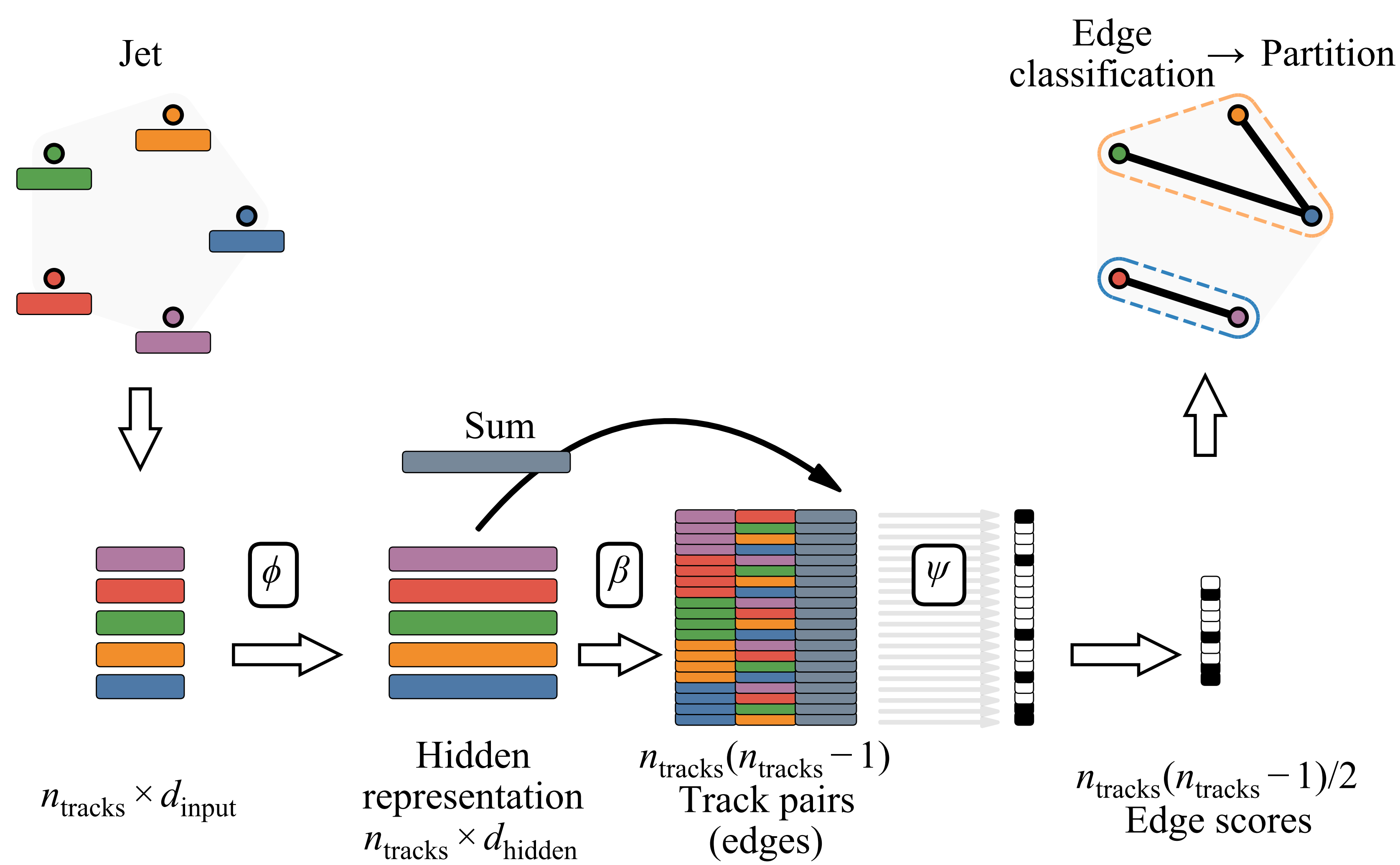



Regular Article - Experimental Physics | Open Access | Published: 23 June 2021

Secondary vertex finding in jets with neural networks

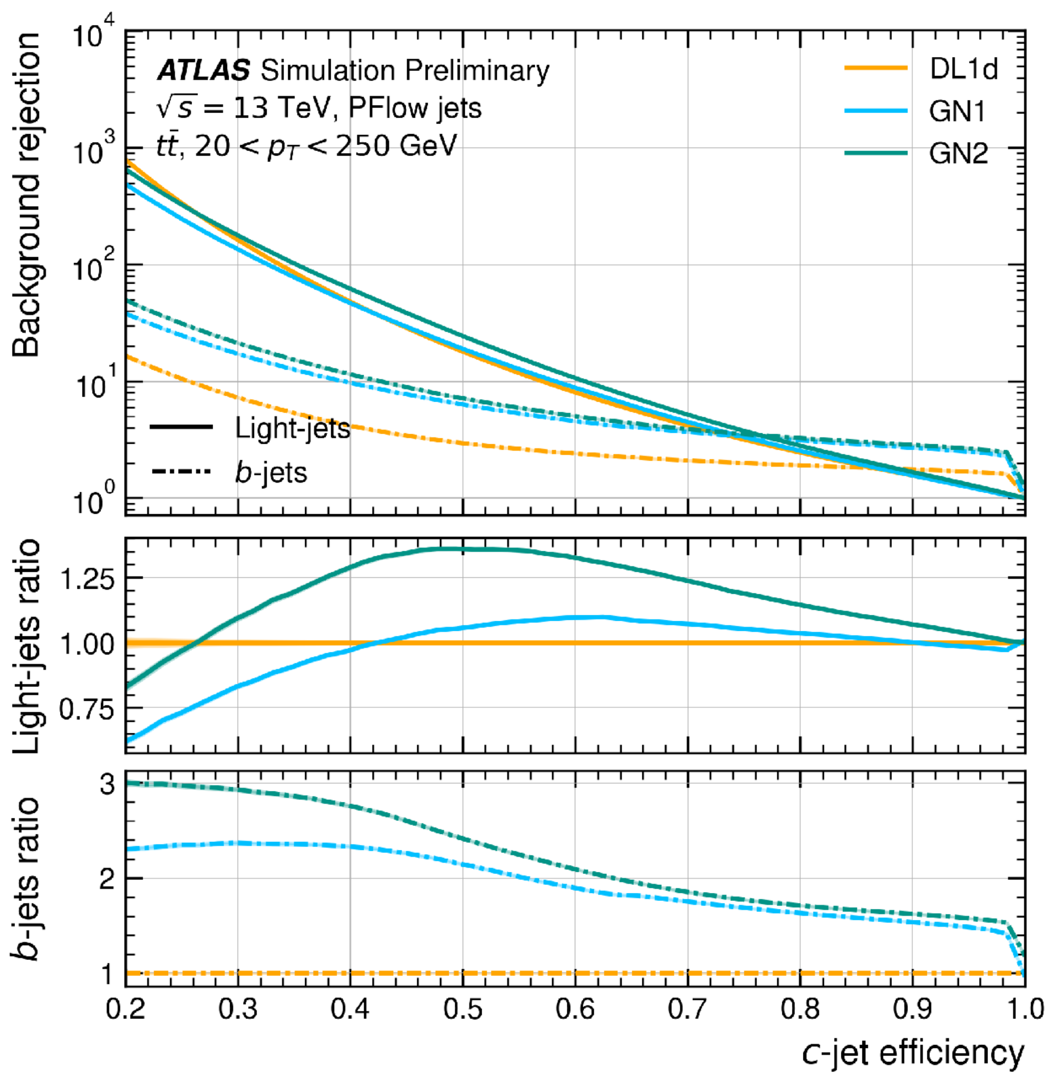
Jonathan Shlomi, Sanmay Ganguly, Eilam Gross, Kyle Cranmer, Yaron Lipman, Hadar Serviansky, Haggai Maron & Nimrod Segol

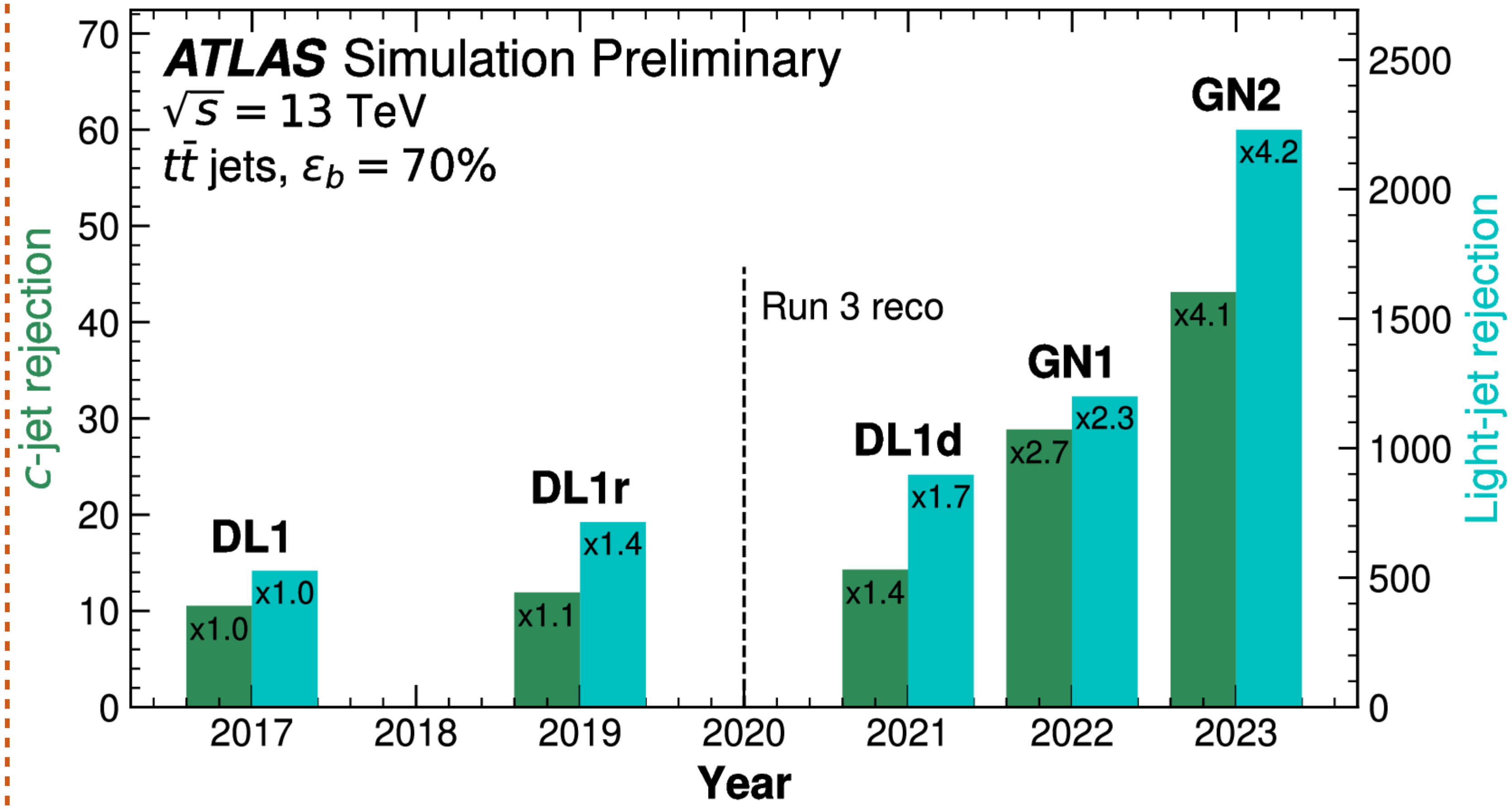
The European Physical Journal C **81**, Article number: 540 (2021) | Cite this article

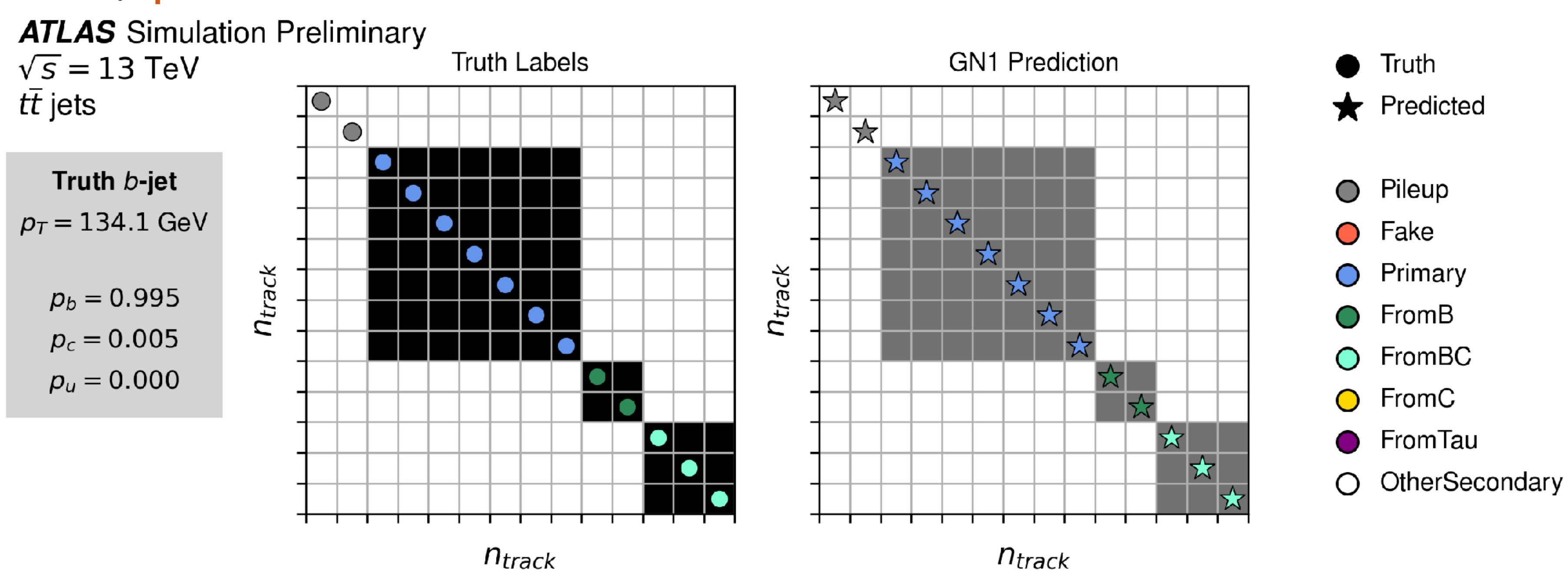


Set2Graph model within ATLAS

FTAG-2023-001





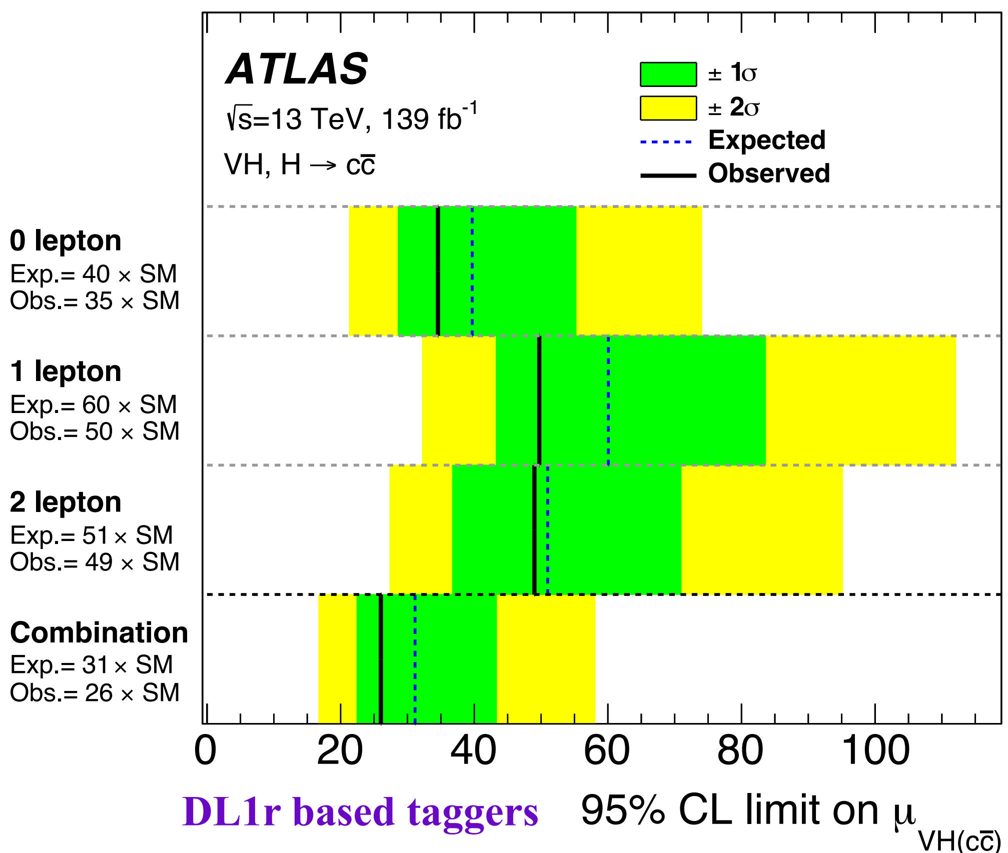


Sizable improvement over the current DL1r algorithm.

For a c-tagging working point ~ 30%, a significant gain in Rejection rate is obtained.

Direct physics application of the taggers

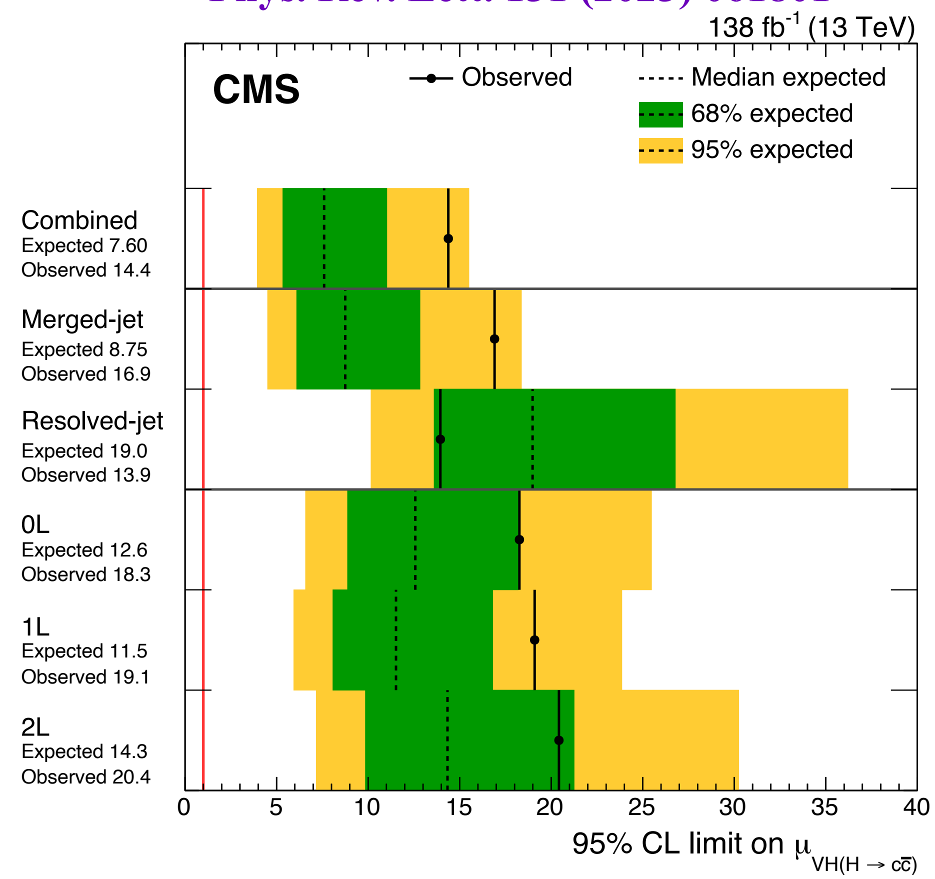
Eur. Phys. J. C (2022) 82:717



DL1r based taggers

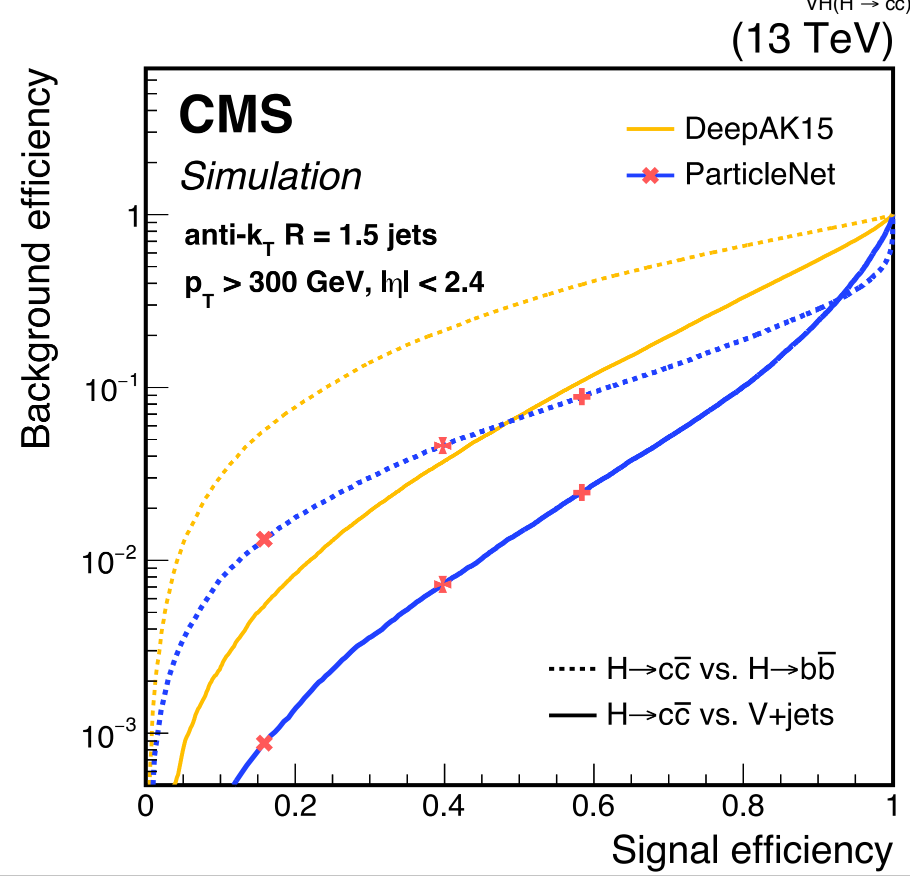
Future iteration will use GN2

Phys. Rev. Lett. 131 (2023) 061801



ATLAS bound : $|\kappa_c| < 8.5$

CMS bound : $1.1 < |\kappa_c| < 5.5$

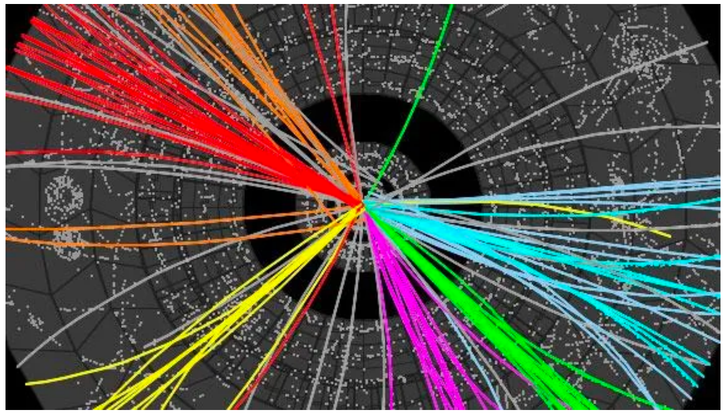


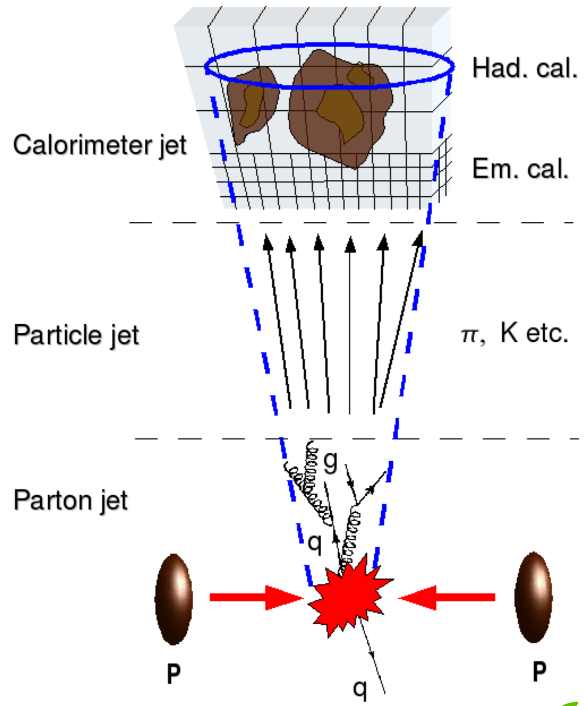
Future direction of tagger improvement:

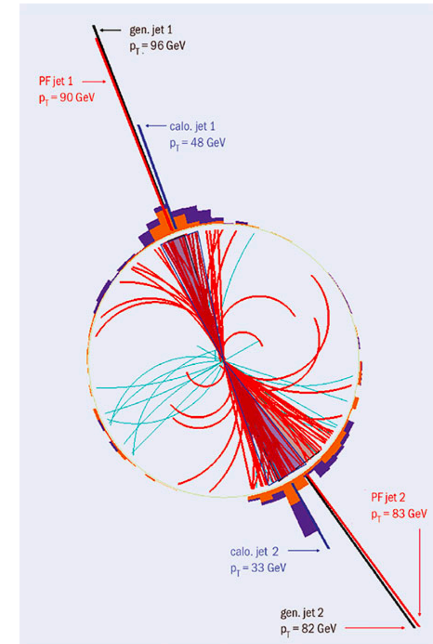
1. **Explainable taggers on heterogeneous pc**
2. **A systematic uncertainty extraction.**
3. **How much universal taggers can be made across topologies?**

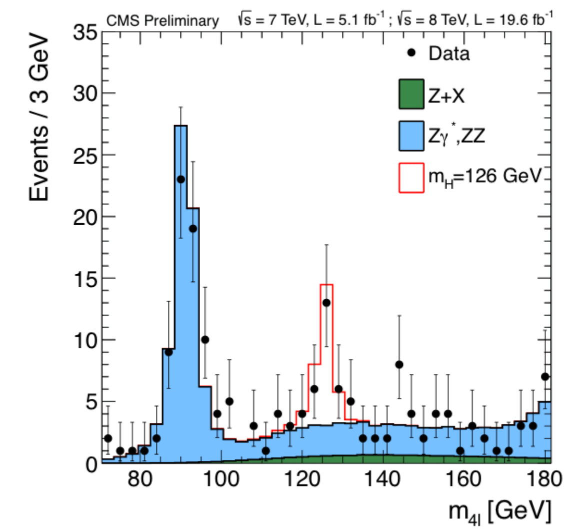
The LHC data flow-chain

ML can play a role at every instance of this flow chain.









✓

Trigger

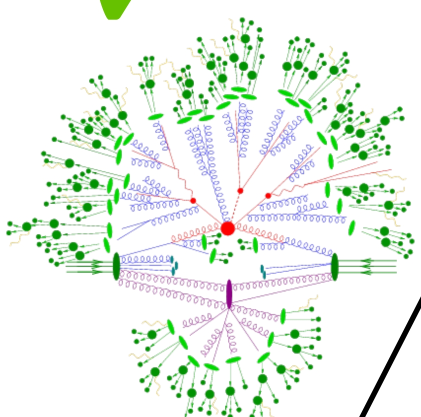
Object rec ✓

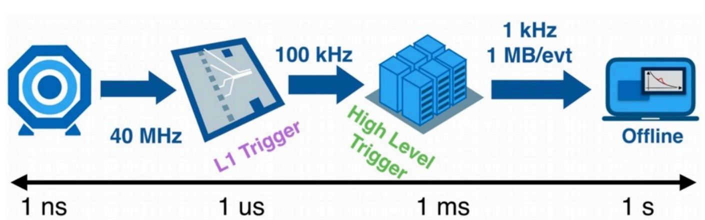
Event rec

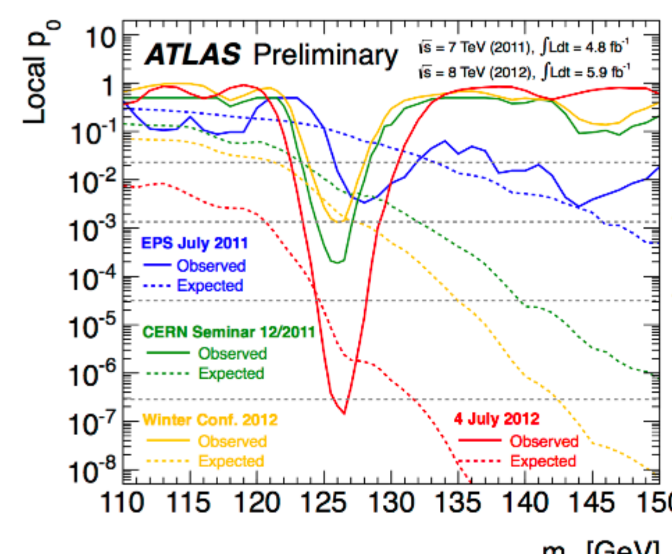
⊙

Analysis

✓

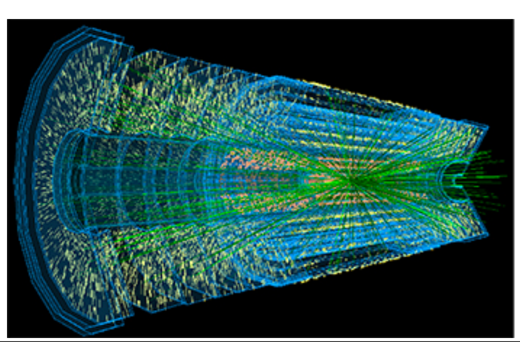






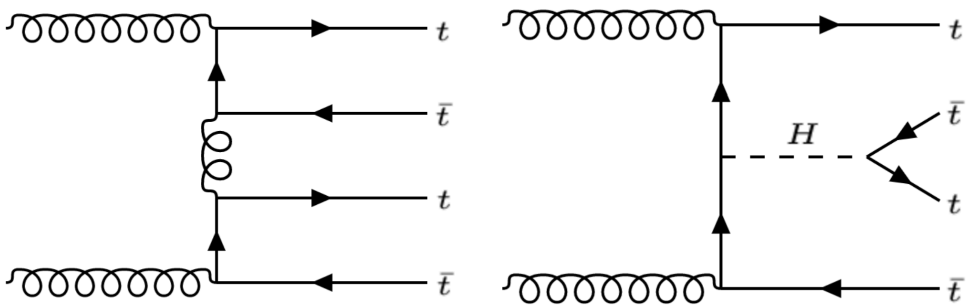
Inference

ME + PS generation
Detector Simulation



✓

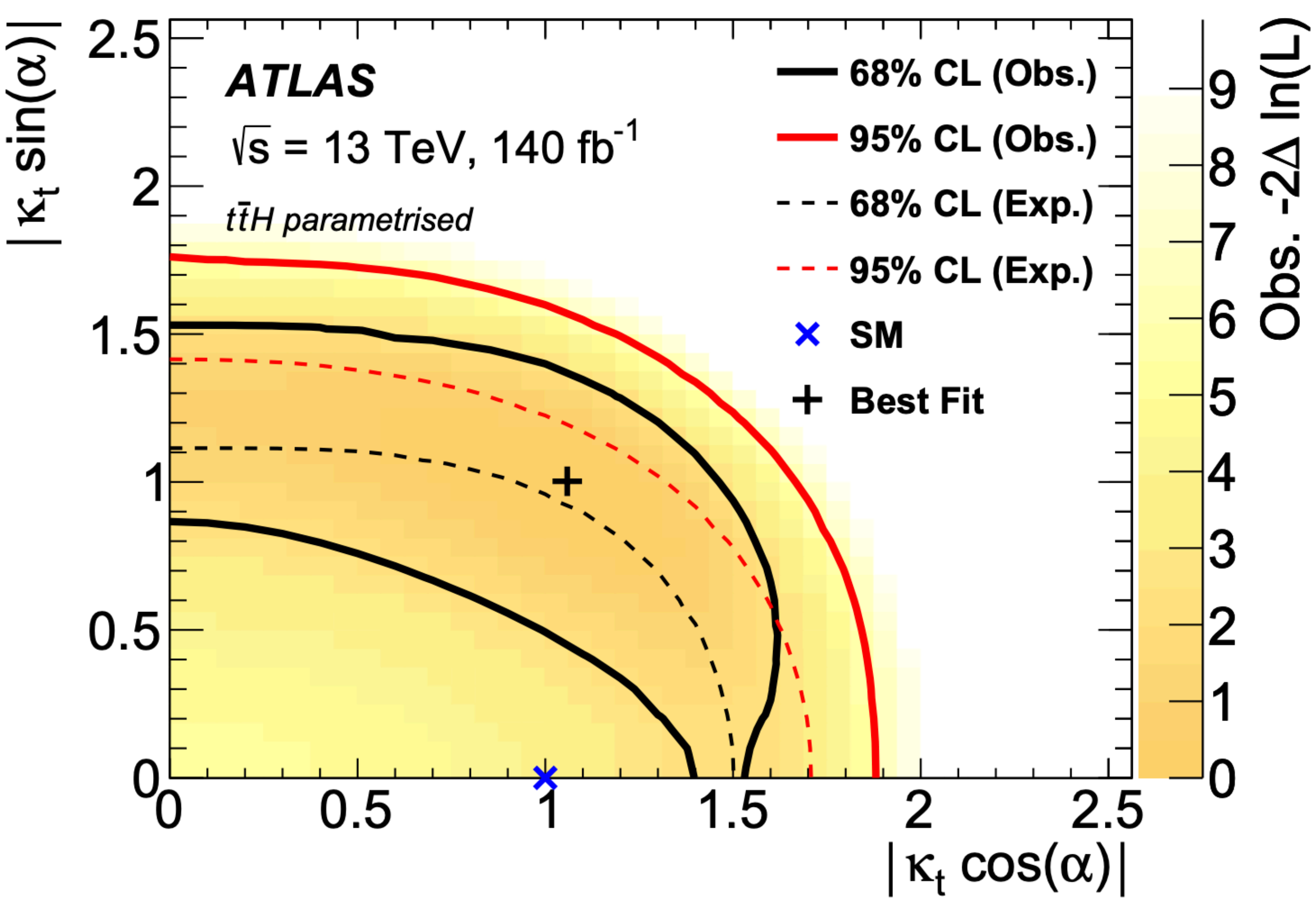
Event construction using NN

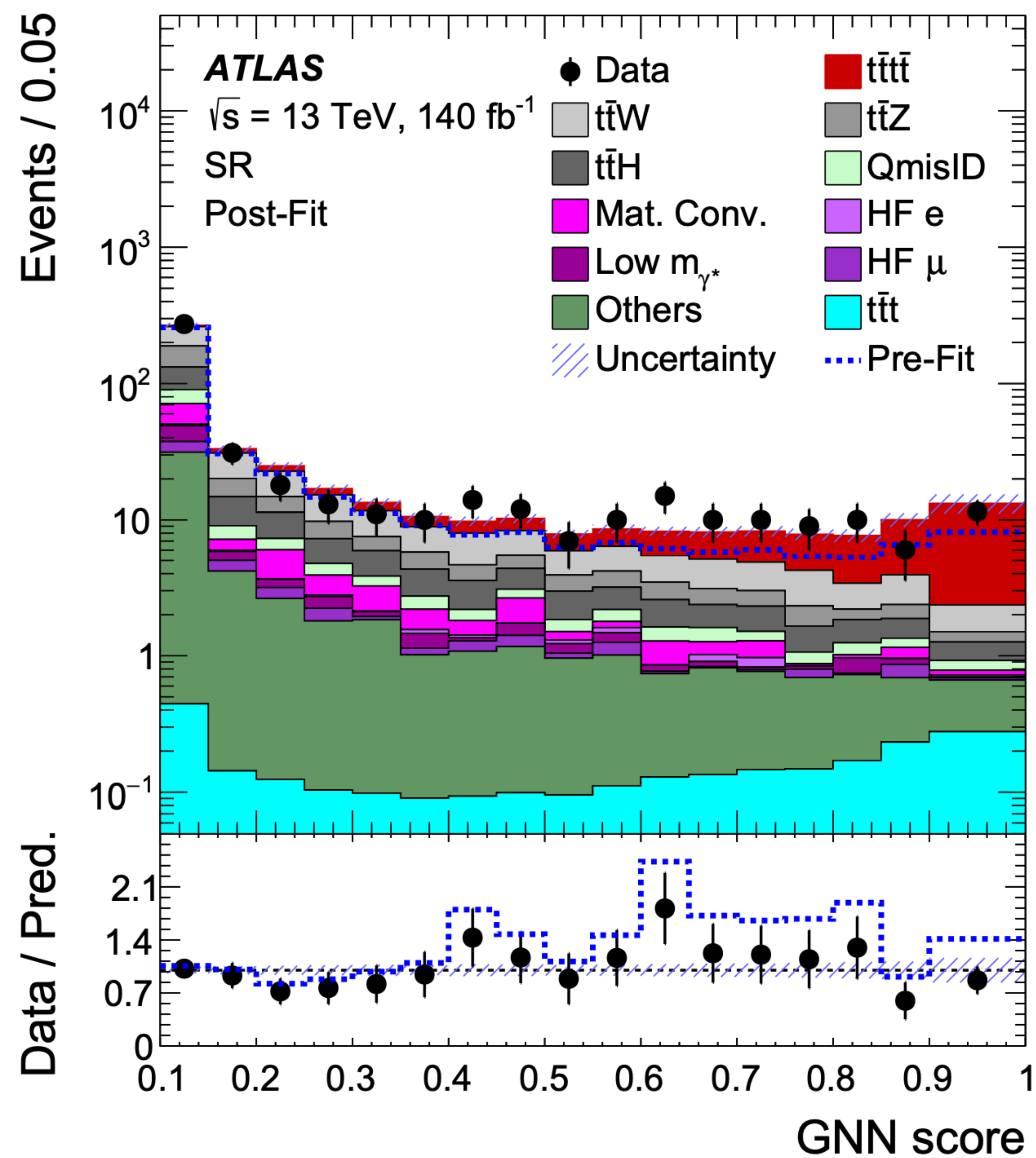


$$\sigma_{t\bar{t}t\bar{t}} = 22.5^{+4.7}_{-4.3}(\text{stat})\,^{+4.6}_{-3.4}(\text{syst})\text{ fb} = 22.5\,^{+6.6}_{-5.5}\text{ fb.}$$

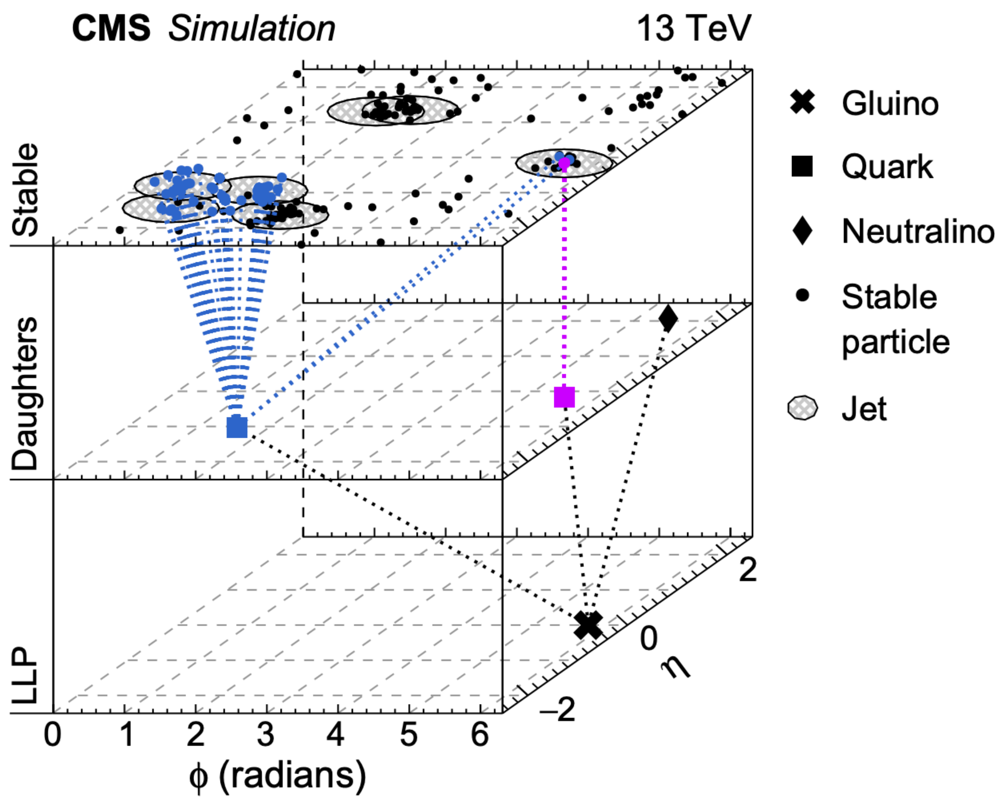
Eur. Phys. J. C 83 (2023) 496

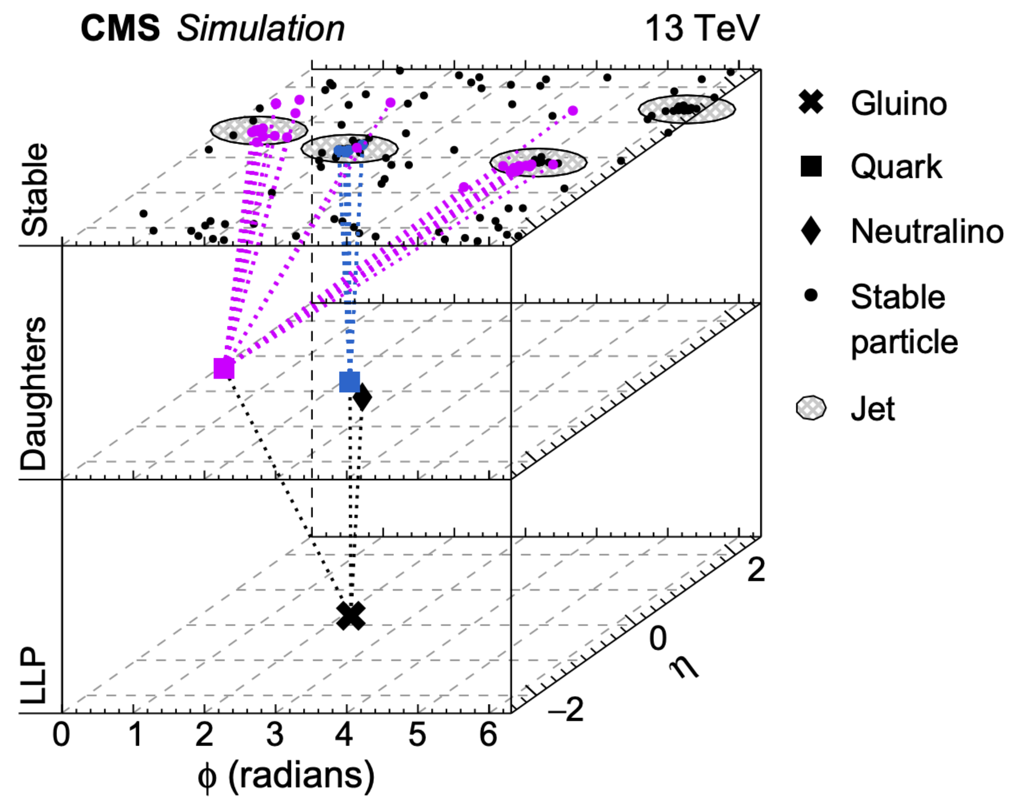
Operators	Expected C_i/Λ^2 [TeV $^{-2}$]	Observed C_i/Λ^2 [TeV $^{-2}$]
O^1_{QQ}	[-2.4, 3.0]	[-3.5, 4.1]
O^1_{Qt}	[-2.5, 2.0]	[-3.5, 3.0]
O^1_{tt}	[-1.1, 1.3]	[-1.7, 1.9]
O^8_{Qt}	[-4.2, 4.8]	[-6.2, 6.9]



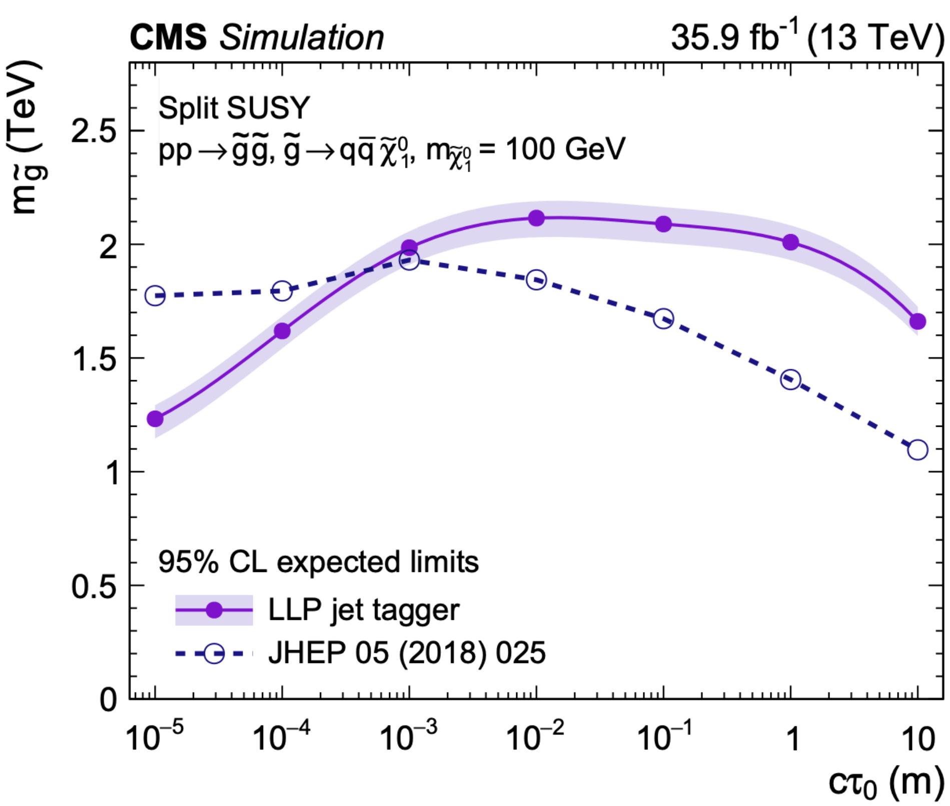


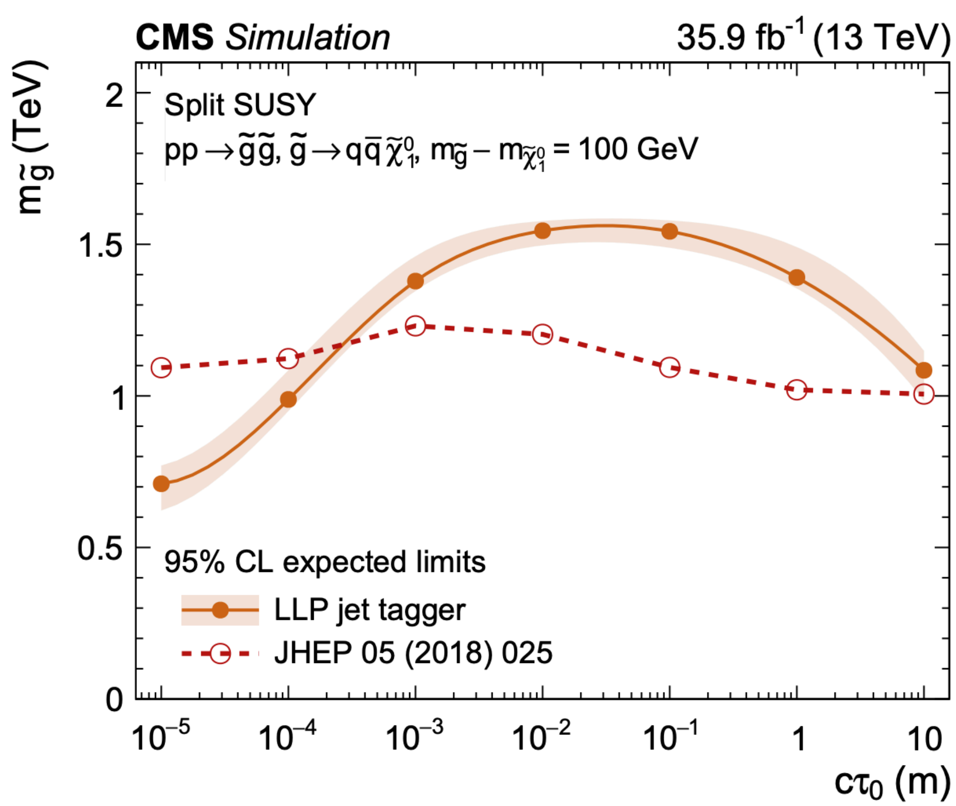
NN's are handy for searching NP





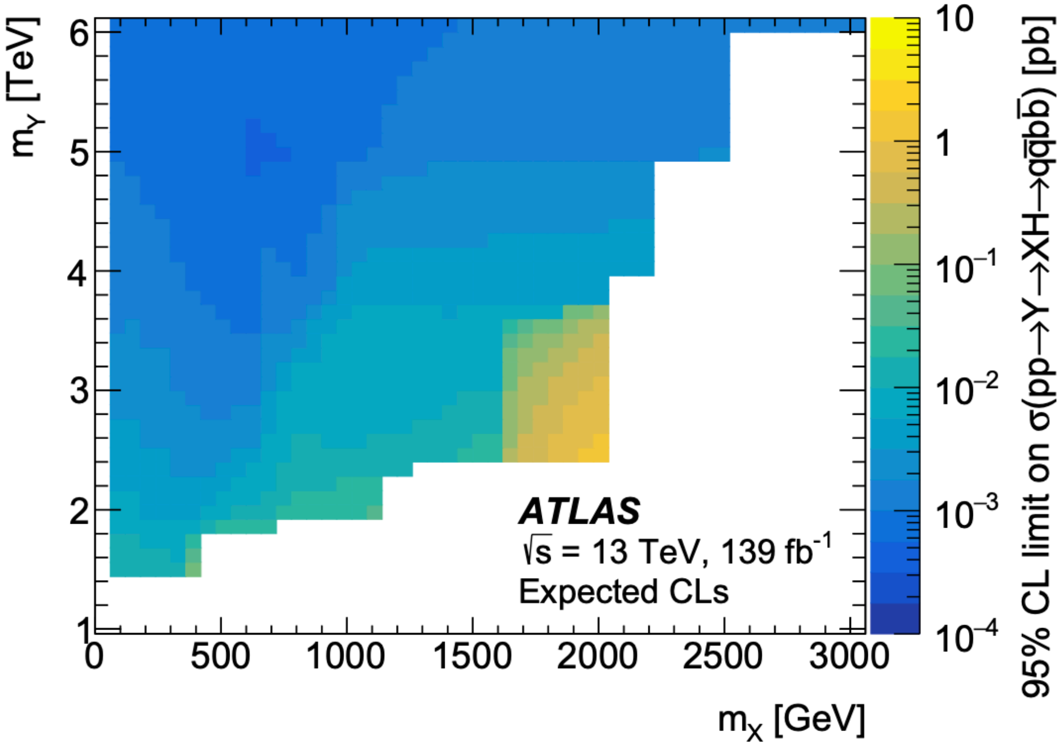
Mach. Learn.: Sci. Technol. 1 (2020) 035012

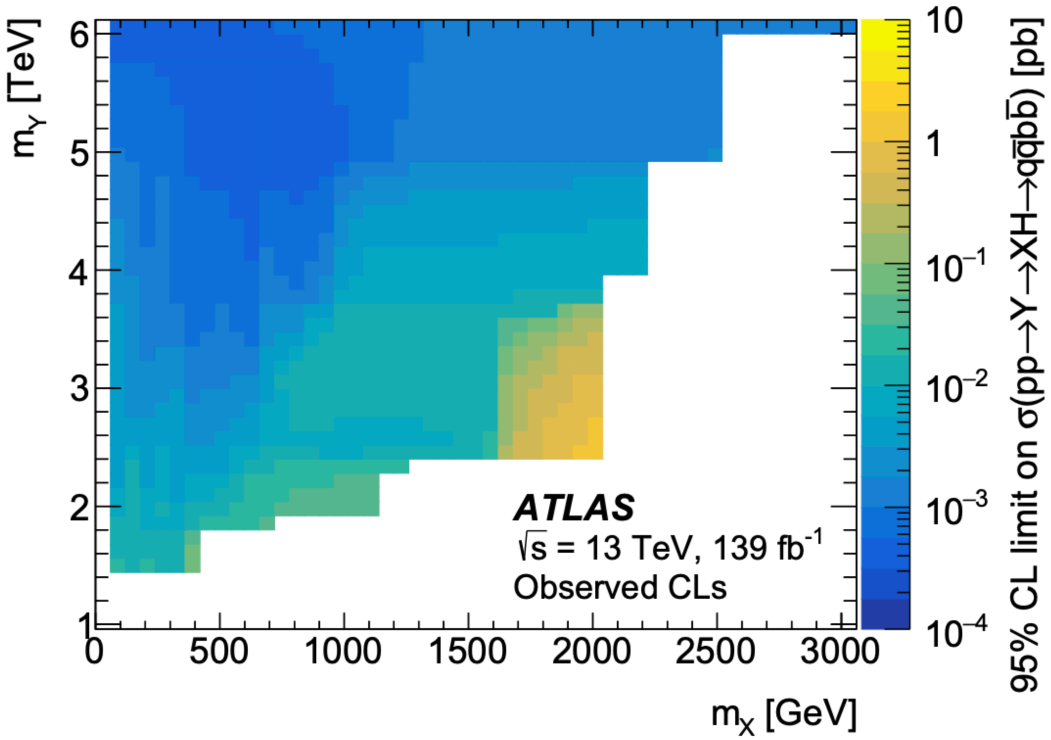


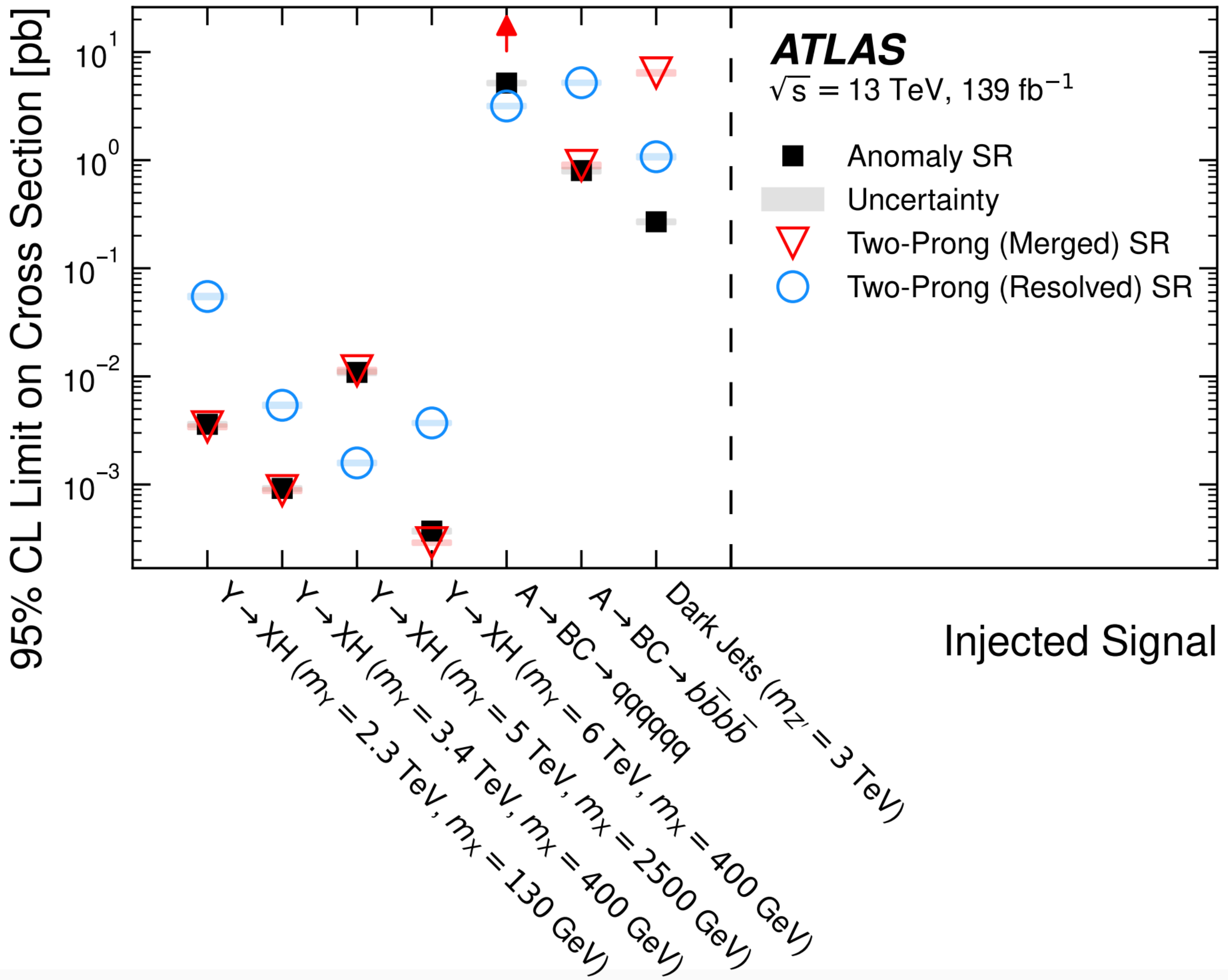


Anomaly detection using NN

$$\mathcal{L}(t) = |\mathbf{y}(t) - \mathbf{x}(t)|^2 + \lambda D_{\text{KL}}(z||z_t)$$



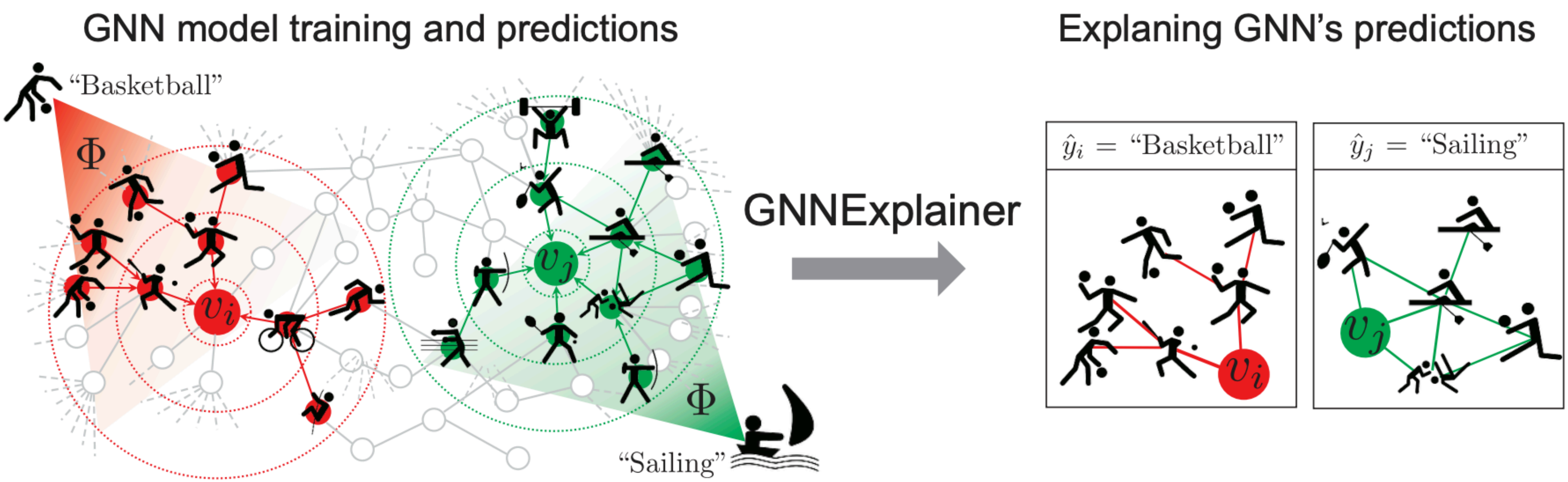




Phys. Rev. D 108 (2023) 052009

Future Directions

Major thrust in immediate future : Interpretability



Interpretability is a key issue and efforts are ongoing to map the NN explainability to first principle physics intuition

Interpretability : an example attempt

$$\mathbf{R}_j^{(l)} = \sum_k \frac{x_j A_{jk}}{\sum_m x_m A_{mk}} \mathbf{R}_k^{(l+1)} \tag{3}$$

where $\mathbf{R}_j^{(l)}$ represent the R-scores of the features of node j at layer l, while the quantity $x_j A_{jk}$ models the extent to which node j at layer l, with activation x_j, contributes to the relevance of node k at layer $l+1$, where A is the adjacency matrix.

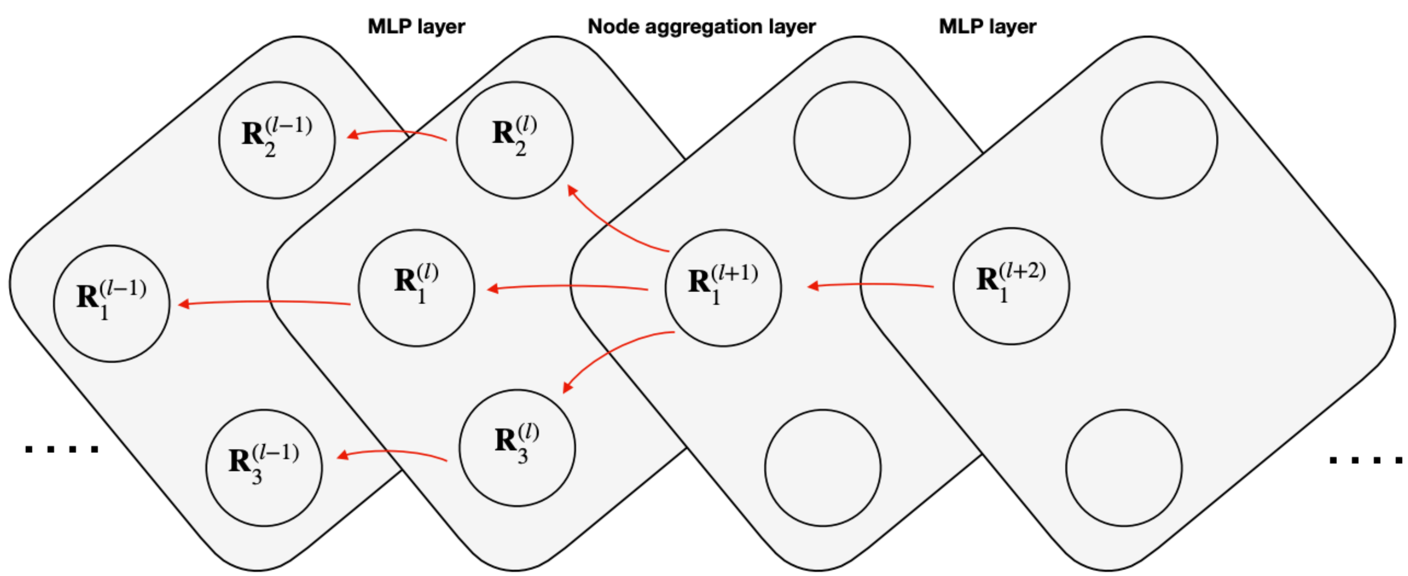
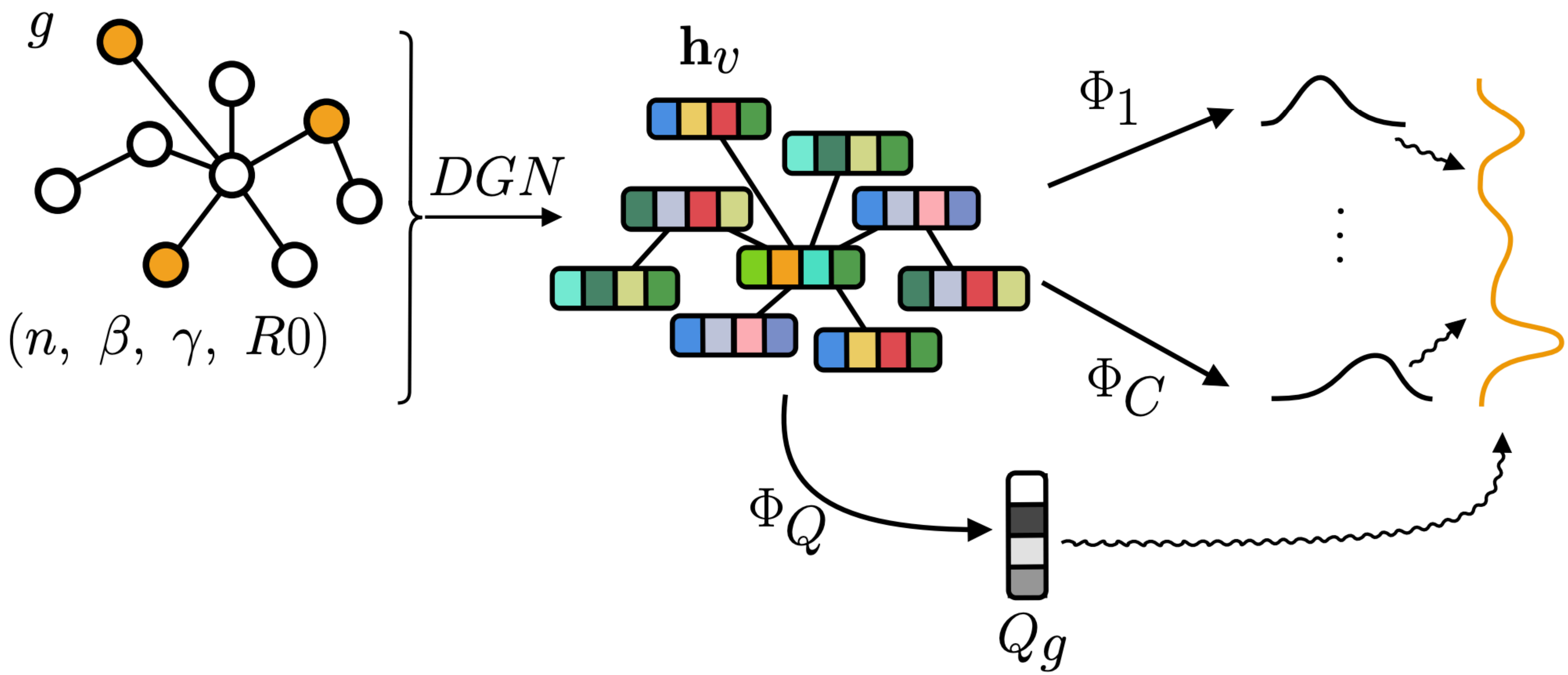


Figure 1: The flow of R-scores of node 1 across the different layers in MLPF. For MLP layers, the redistribution of R-scores follows the standard LRP rules [35, 36]. For the aggregation step in the message passing layer, the redistribution follows Equation 3. We only show three nodes for simplicity.

Explainability for MLPF

Neur IPS 2021. F. Mokhtar, R. Kansal et al

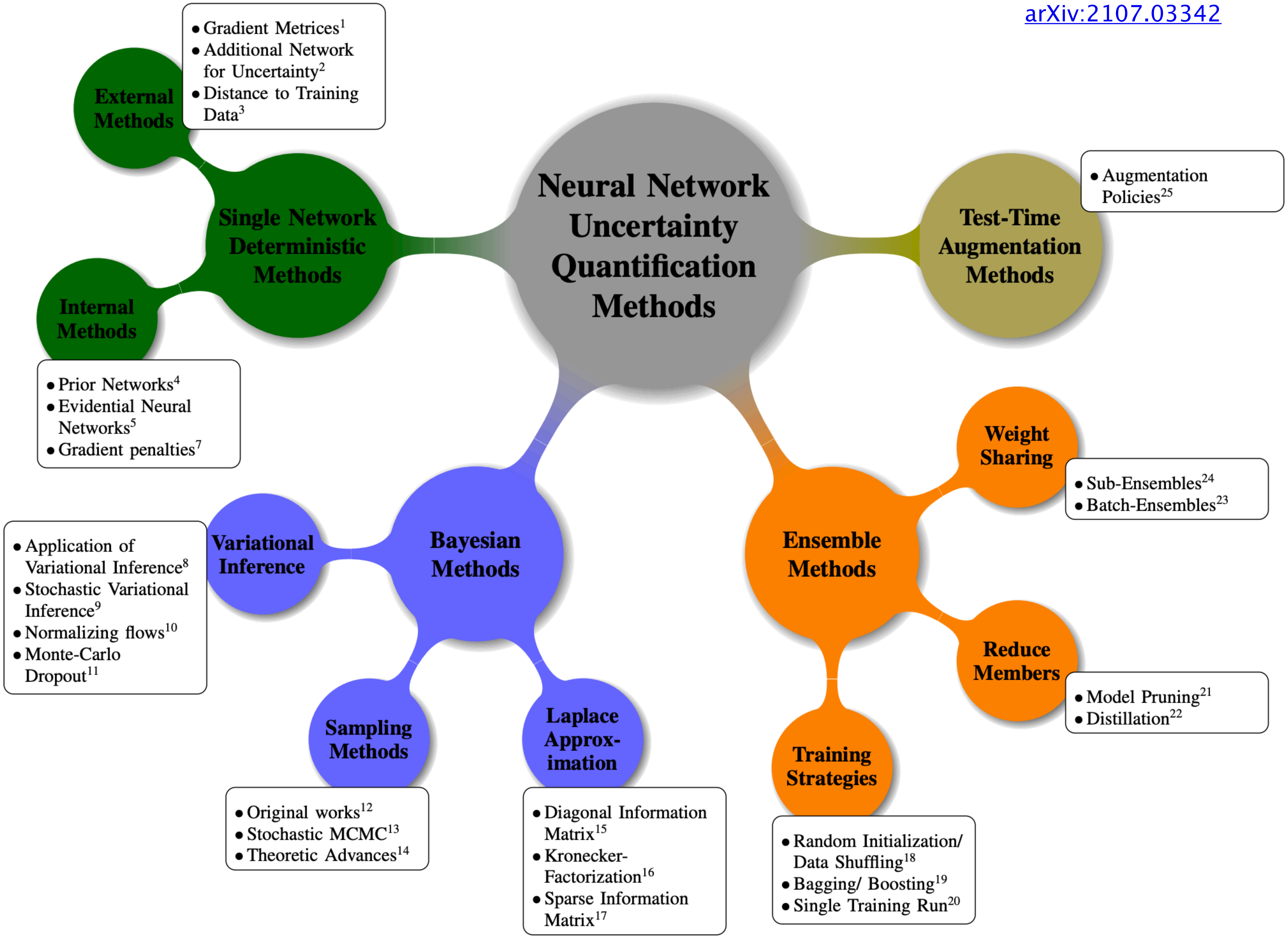
Major thrust in immediate future : Uncertainty



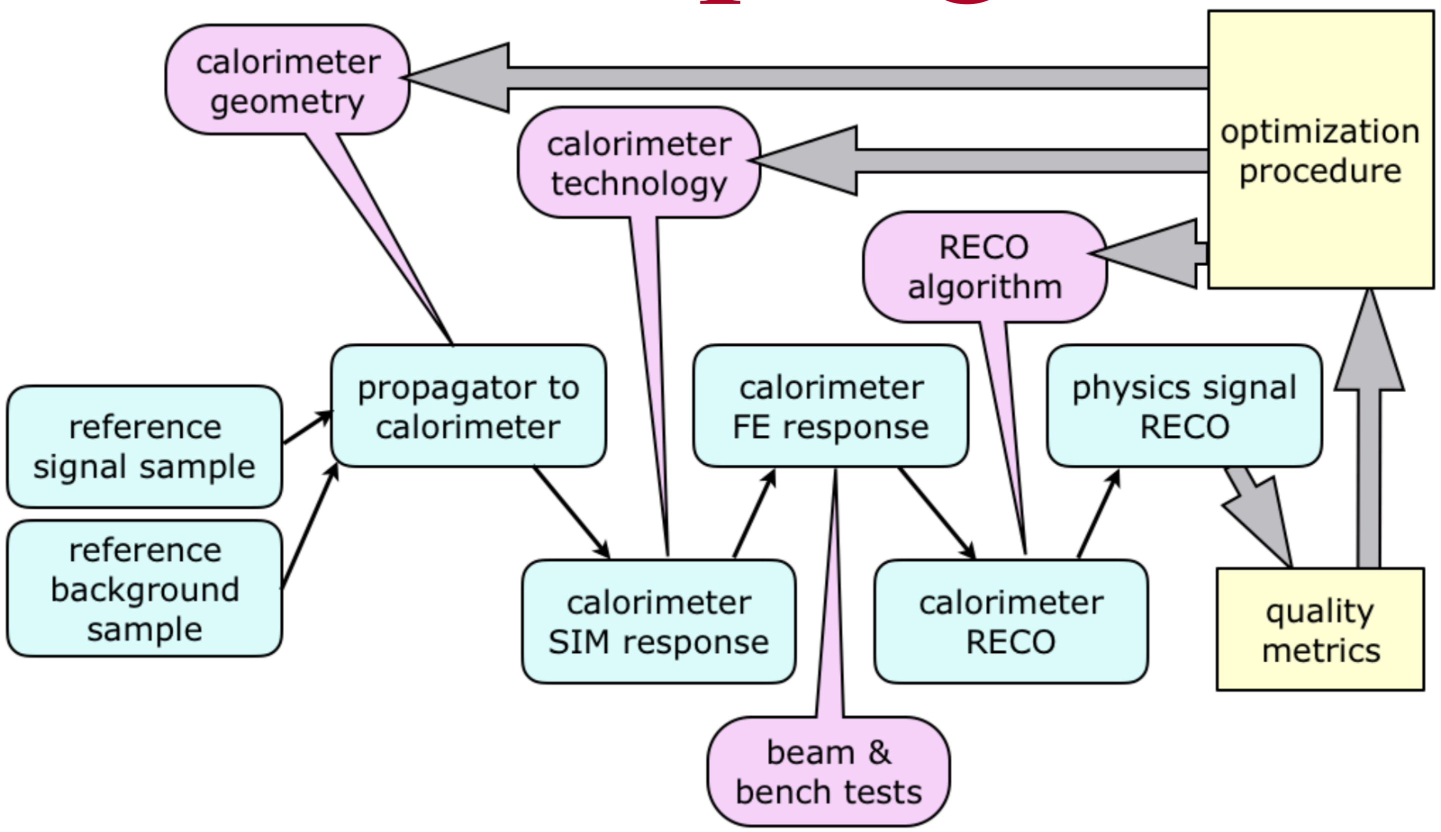
Reliable uncertainty estimation on ML based predictions are crucial for HEP
Only few Bayesian methods have been tested naively.

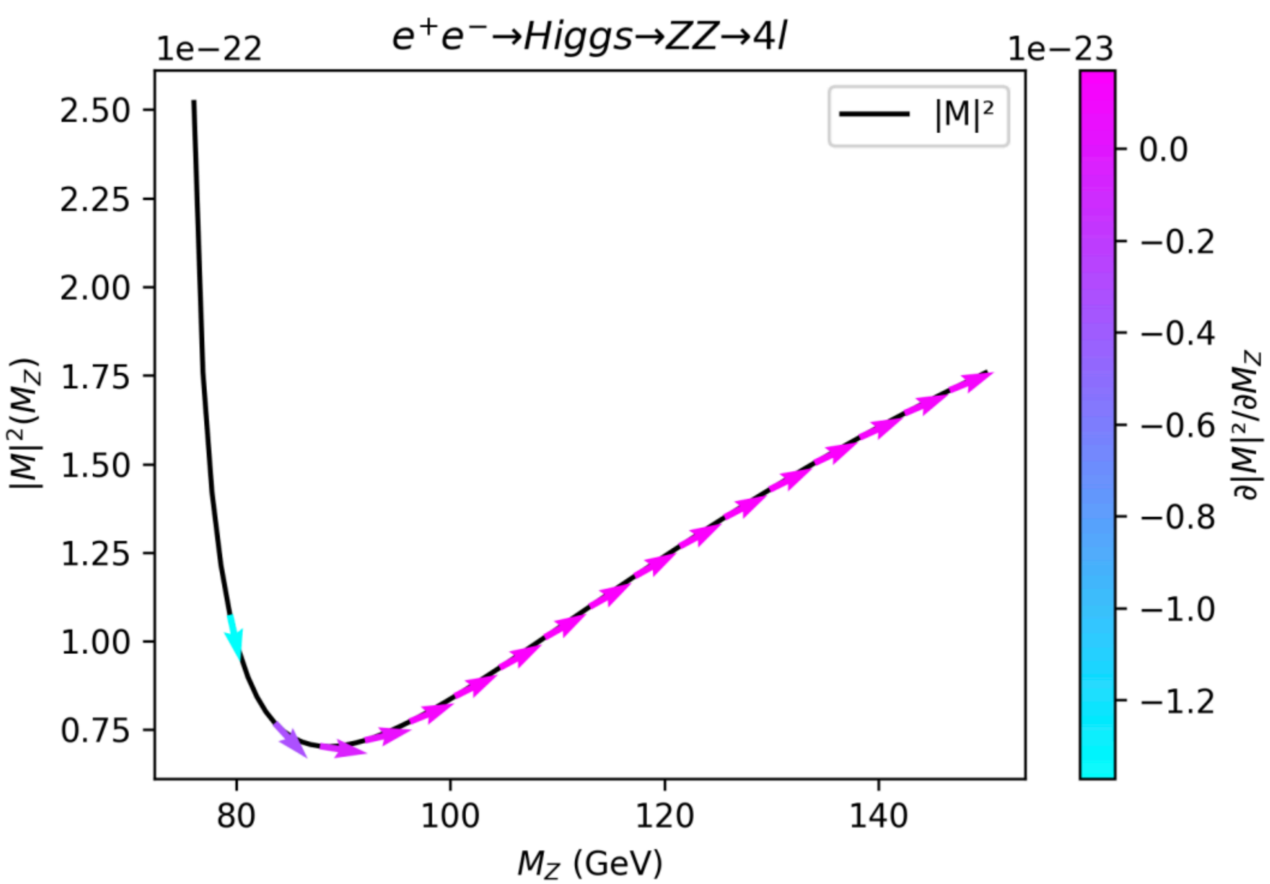
Major thrust in immediate future : Uncertainty

arXiv:2107.03342



Differential programming in HEP





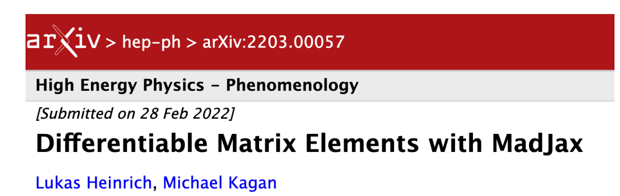

```
generate p p > t t~, t > b udsc udscx , t~ > b~ udsc udscx
output madjax generated_ttbar
set auto_update 0
```

2. Evaluation:

```
import madjax
mj = madjax.MadJax('generated_ttbar')
E_cm = 14000 #GeV
process = 'Matrix_1_gg_ttx_t_budx_tx_bxdux'
matrix_element = mj.matrix_element(E_cm,process)

parameters = ('mass',6): 173.0 #set top mass
phasespace_coords = [0.1]*14 #14D phasespace

val, grad = matrix_element(parameters,phasespace_coords)
grad[('mass', 6)] #gradient wrt top mass
```


arXiv > hep-ph > arXiv:2203.00057

High Energy Physics – Phenomenology

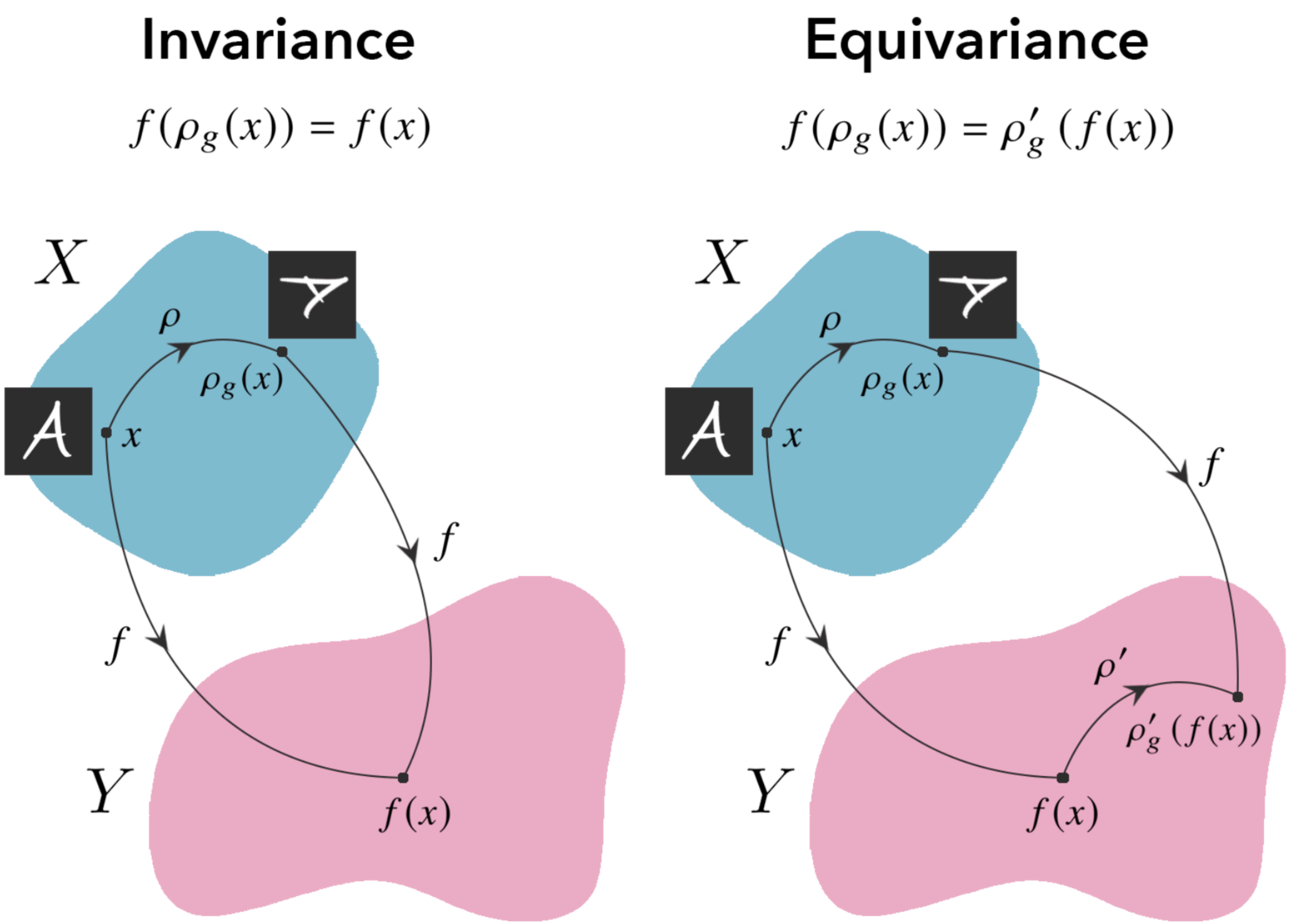
[Submitted on 28 Feb 2022]

Differentiable Matrix Elements with MadJax

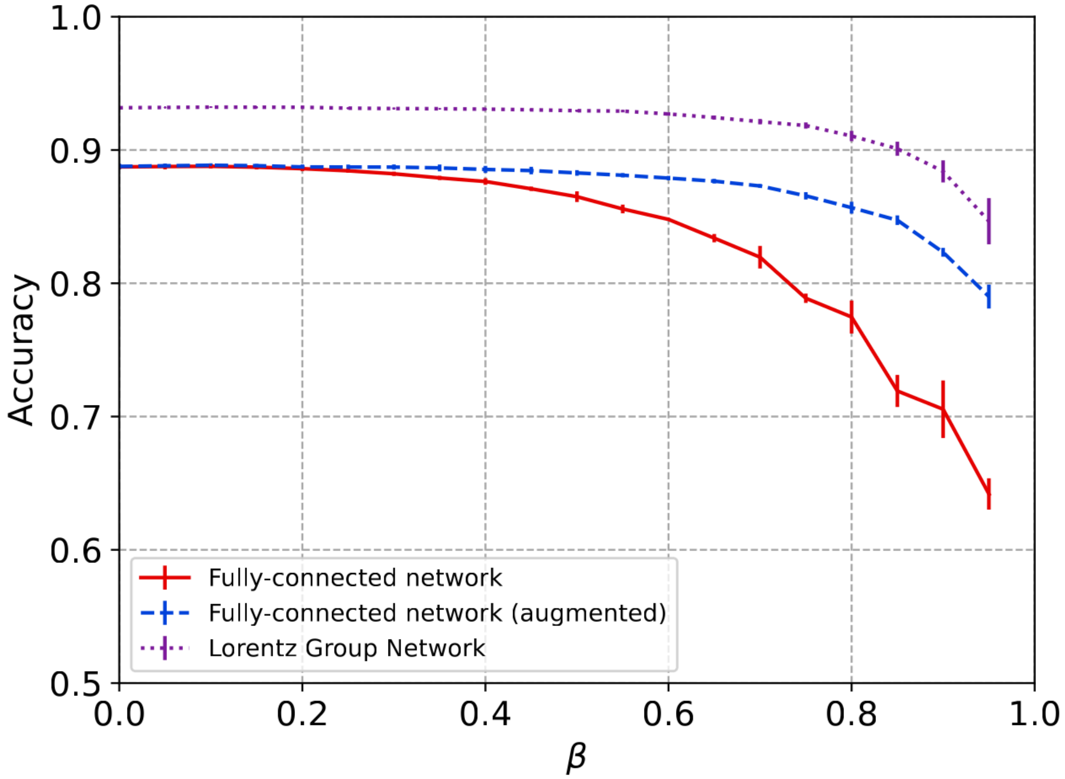
Lukas Heinrich, Michael Kagan

Symmetry equivariant networks

arXiv:2203.06153 : SG et al



Symmetry equivariant networks

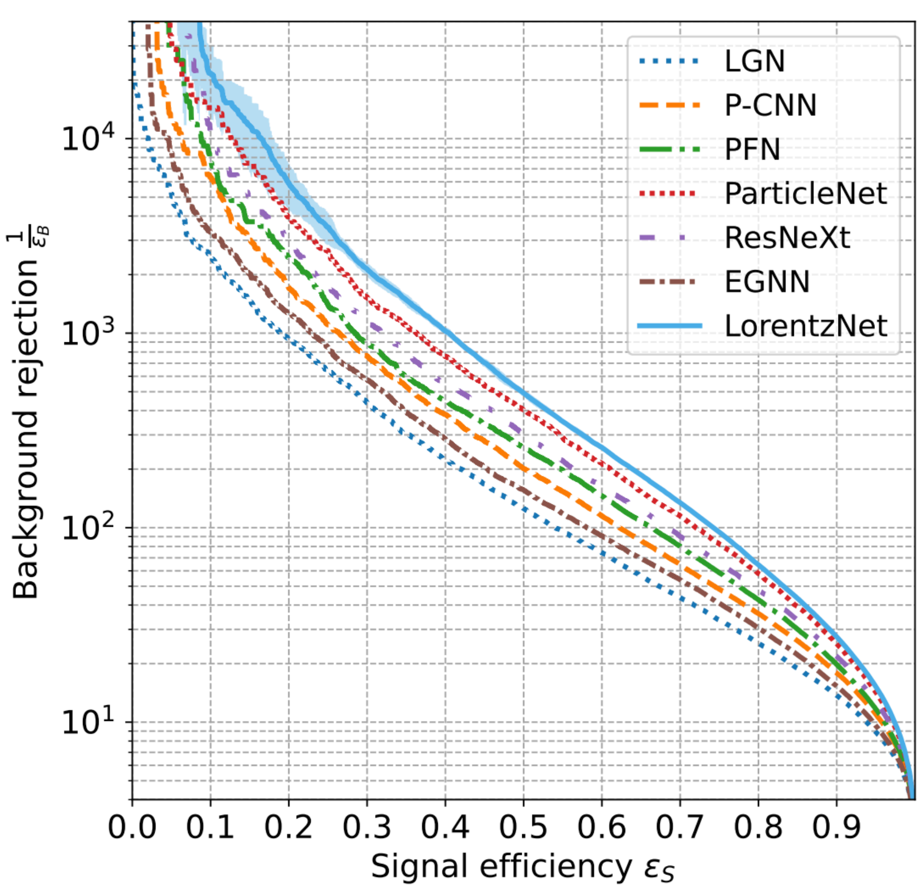


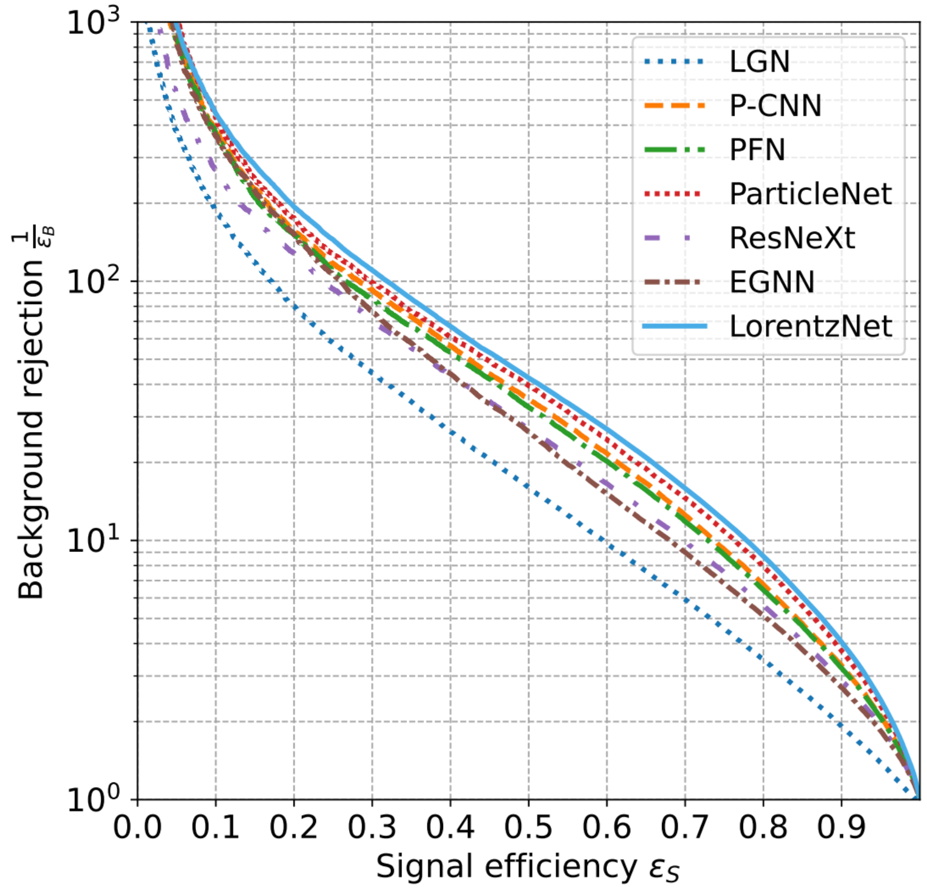
arXiv:2006.04780 : A Bogatskiy et al

arXiv:2203.06153 : SG et al

$$m_{ij}^l = \phi_e\left(h_i^l, h_j^l, \psi(\|x_i^l - x_j^l\|^2), \psi(\langle x_i^l, x_j^l \rangle)\right)$$

$$h_i^{l+1} = h_i^l + \phi_h(h_i^l, \sum_{j\in[N]} w_{ij} m_{ij}^l),$$





LG equivariant GNN : arXiv 2201.08187

Field theory & ML

arXiv > hep-th > arXiv:2202.11737

Search...

Help | Advan

High Energy Physics – Theory

[Submitted on 23 Feb 2022 (v1), last revised 14 Mar 2022 (this version, v2)]

Renormalization Group Flow as Optimal Transport

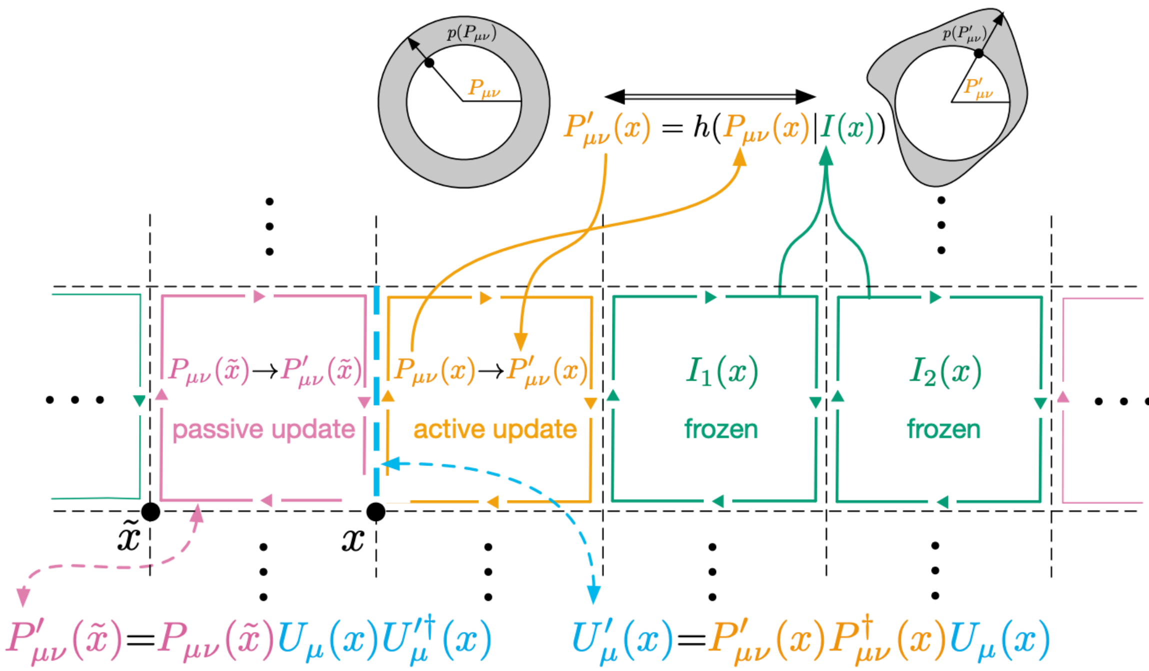
Jordan Cotler, Semon Rezchikov

We establish that Polchinski's equation for exact renormalization group flow is equivalent to the optimal transport gradient flow of a field-theoretic relative entropy. This provides a compelling information-theoretic formulation of the exact renormalization group, expressed in the language of optimal transport. A striking consequence is that a regularization of the relative entropy is in fact an RG monotone. We compute this monotone in several examples. Our results apply more broadly to other exact renormalization group flow equations, including widely used specializations of Wegner-Morris flow. Moreover, our optimal transport framework for RG allows us to reformulate RG flow as a variational problem. This enables new numerical techniques and establishes a systematic connection between neural network methods and RG flows of conventional field theories.

$$
\begin{aligned}
&- \int [d\phi]\, \Lambda \frac{\partial P_\Lambda^{\text{free}}[\phi]}{\partial \Lambda}\, S_{\text{free},\Lambda}[\phi] \\
&\quad = \frac{1}{4 Z_{P,\Lambda}} \int [d\phi]\, e^{-\frac{1}{2}\int \frac{d^d p}{(2\pi)^d}\, \phi(p)\phi(-p)(p^2+m^2)K_\Lambda^{-1}(p^2)} \int \frac{d^d p}{(2\pi)^d}\, \phi(p)\phi(-p)(p^2+m^2)K_\Lambda^{-1}(p^2) \\
&\qquad\qquad \times \int \frac{d^d q}{(2\pi)^d}\, \phi(q)\phi(-q)(q^2+m^2)\Lambda \frac{\partial K_\Lambda^{-1}(q^2)}{\partial \Lambda} \\
&\qquad + \Lambda \frac{\partial \log Z_{P,\Lambda}}{\partial \Lambda} \int [d\phi]\, P_\Lambda^{\text{free}}[\phi]\, S_{\text{free},\Lambda}[\phi] \\
&\quad = \left(\frac{1}{2}\int d^d p\, \delta^d(0)\right)\left(\Lambda \frac{\partial \log Z_{P,\Lambda}}{\partial \Lambda} - \frac{1}{2}\delta^d(0)\int d^d p\, \Lambda \frac{\partial \log K_\Lambda(p^2)}{\partial \Lambda}\right) - \frac{1}{2}\delta^d(0)\int d^d p\, \Lambda \frac{\partial \log K_\Lambda(p^2)}{\partial \Lambda}\,.
\end{aligned}
$$

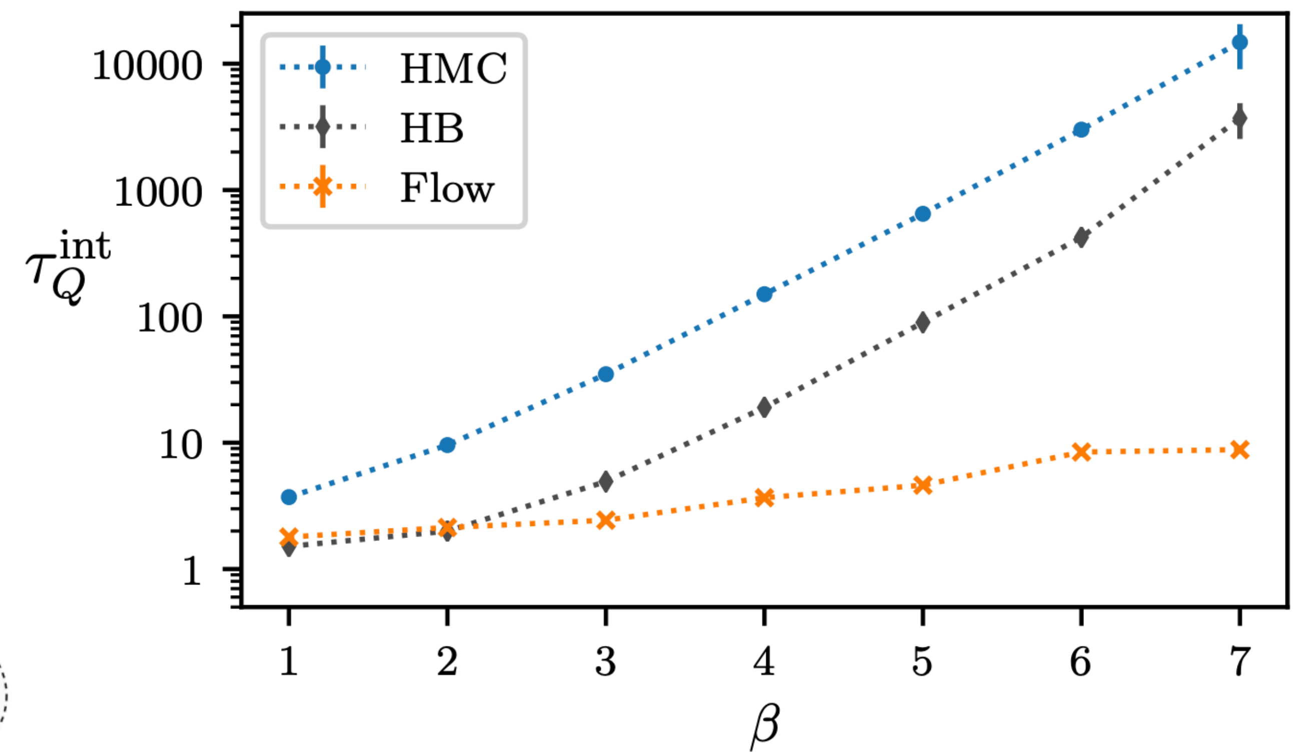
$$
\begin{aligned}
&\mathsf{M}_{\widetilde{\mathcal{F}}_{n\tau *}P_{\Lambda_0}}[\mathcal{F}] \\
&= \frac{1}{2}\int [d\phi]\, P_{\Lambda_0}[\phi] \int d^d x\, d^d y\, \dot{C}_\Lambda(x,y)\Big(\widetilde{\mathcal{F}}_{n\tau}[\phi(x)] - (\mathcal{F}\circ\widetilde{\mathcal{F}}_{n\tau})[\phi(x)]\Big)\Big(\widetilde{\mathcal{F}}_{n\tau}[\phi(y)] - (\mathcal{F}\circ\widetilde{\mathcal{F}}_{n\tau})[\phi(y)]\Big) \\
&= \mathop{\mathbb{E}}_{P_{\Lambda_0}}\left[\frac{1}{2}\int d^d x\, d^d y\, \dot{C}_\Lambda(x,y)\Big(\widetilde{\mathcal{F}}_{n\tau}[\phi(x)] - (\mathcal{F}\circ\widetilde{\mathcal{F}}_{n\tau})[\phi(x)]\Big)\Big(\widetilde{\mathcal{F}}_{n\tau}[\phi(y)] - (\mathcal{F}\circ\widetilde{\mathcal{F}}_{n\tau})[\phi(y)]\Big)\right].
\end{aligned}
$$

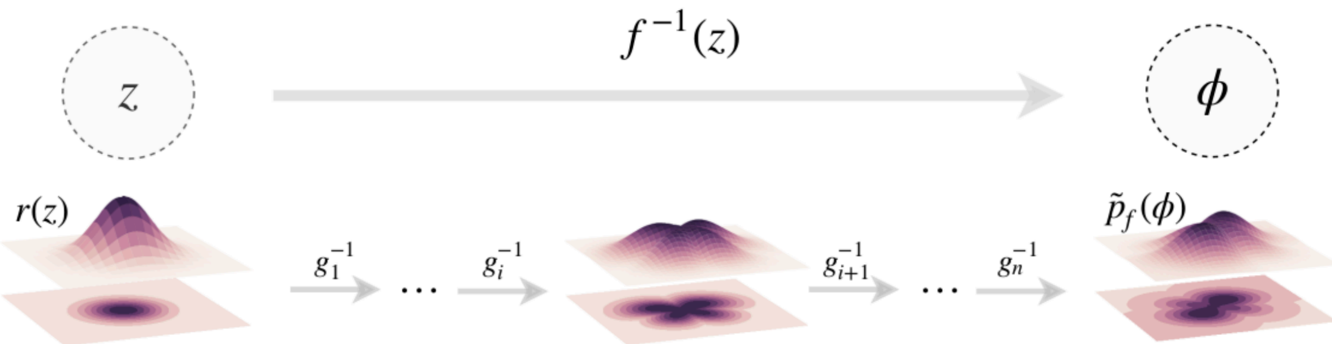
Lattice field theory & ML



When model parameters are tuned towards criticality, the Traditional MCMC based sampling slows down and there Gauge Equivariant Normalizing Flow based methods are more handy

G. Kanwar et al : *Phys.Rev.Lett.* 125 (2020) 12, 121601





arXiv : 2101.08176

Applications of Machine Learning to Lattice Quantum Field Theory

Denis Boyda (Argonne, ALCF and IAIFI, Cambridge), Salvatore Calì (MIT, Cambridge, CTP and IAIFI, Cambridge), Sam Foreman (Argonne, ALCF), Lena Funcke (MIT, Cambridge, CTP and IAIFI, Cambridge and Unlisted), Daniel C. Hackett (MIT, Cambridge, CTP and IAIFI, Cambridge) Show All(11)

Feb 10, 2022

10 pages

Contribution to: 2022 Snowmass Summer Study

e-Print: 2202.05838 [hep-lat]

Report number: MIT-CTP/5405

View in: ADS Abstract Service

Lattice field theory & ML

arXiv > physics > arXiv:2308.08615

Physics > Computational Physics

[Submitted on 16 Aug 2023 (v1), last revised 23 Dec 2023 (this version, v3)]

Scalable Lattice Sampling using Factorized Generative Models

Ali Faraz, Ankur Singha, Dipankar Chakrabarti, Vipul Arora

arXiv > hep-lat > arXiv:2306.00581

High Energy Physics – Lattice

[Submitted on 1 Jun 2023 (v1), last revised 31 Oct 2023 (this version, v2)]

Sampling U(1) gauge theory using a re-trainable conditional flow-based model

Ankur Singha, Dipankar Chakrabarti, Vipul Arora

If you want to know more : there's a hidden chamber beyond secretariat's office !!

Many branches of science using same stuff

arXiv > astro-ph > arXiv:2111.08683

Astrophysics > Cosmology and Nongalactic Astrophysics

[Submitted on 16 Nov 2021]

Inferring halo masses with Graph Neural Networks

Pablo Villanueva-Domingo, Francisco Villaescusa-Navarro, Daniel Anglés-Alcázar, Shy Genel, Federico Marinacci, David N. Spergel, Lars Hernquist, Mark Vogelsberger, Romeel Dave, Desika Narayanan

Understanding the halo-galaxy connection is fundamental in order to improve our knowledge on the nature and p
model that infers the mass of a halo given the positions, velocities, stellar masses, and radii of the galaxies it host
correlations among galaxy properties and their phase-space, we use Graph Neural Networks (GNNs), that are desi
train our models on galaxies from more than 2,000 state-of-the-art simulations from the Cosmology and Astropl
(CAMELS) project. Our model, that accounts for cosmological and astrophysical uncertainties, is able to constrain t
Furthermore, a GNN trained on a suite of simulations is able to preserve part of its accuracy when tested on simul
distinct subgrid physics model, showing the robustness of our method. The PyTorch Geometric implementation of
https URL

Orbital graph convolutional neural network for material property prediction

Mohammadreza Karamad, Rishikesh Magar, Yuting Shi, Samira Siahrostami, Ian D. Gates, and Amir Barati Farimani

Phys. Rev. Materials **4**, 093801 – Published 8 September 2020

arXiv > cs > arXiv:2203.17003

Computer Science > Machine Learning

[Submitted on 31 Mar 2022]

Equivariant Diffusion for Molecule Generation in 3D

Emiel Hoogeboom, Victor Garcia Satorras, Clément Vignac, Max Welling

This work introduces a diffusion model for molecule generation in 3D that is equivariant to Euclidean transformations. Our E(3) Equivariant Diffusion Model (EDM) learns to denoise a diffusion process with an equivariant network that jointly operates on both continuous (atom coordinates) and categorical features (atom types). In addition, we provide a probabilistic analysis which admits likelihood computation of molecules using our model. Experimentally, the proposed method significantly outperforms previous 3D molecular generative methods regarding the quality of generated samples and efficiency at training time.

Summarized all the important aspects learned and where to keep an eye :

arXiv > hep-ex > arXiv:2203.12852

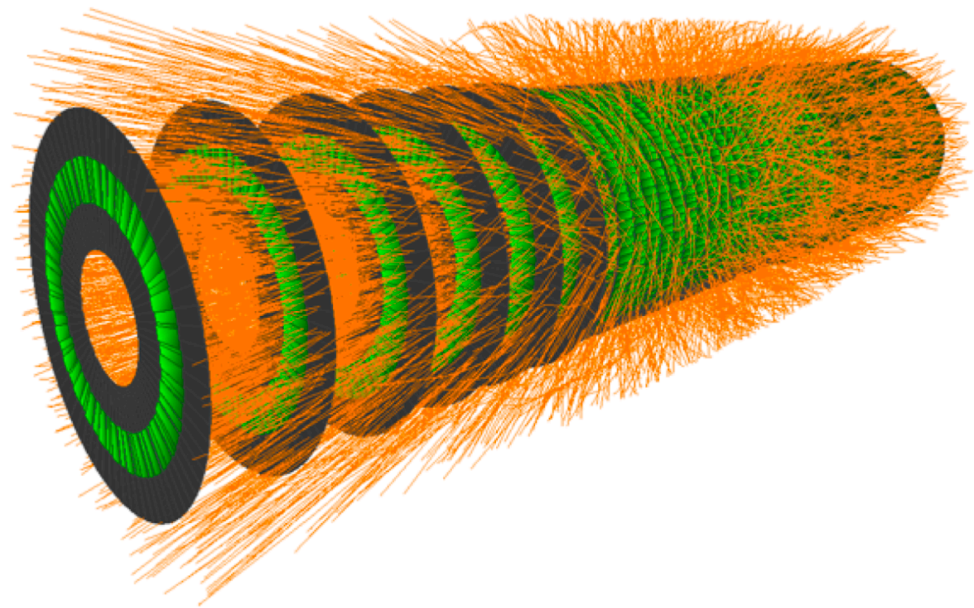
High Energy Physics – Experiment

[Submitted on 23 Mar 2022 (v1), last revised 25 Mar 2022 (this version, v2)]

Graph Neural Networks in Particle Physics: Implementations, Innovations, and Challenges

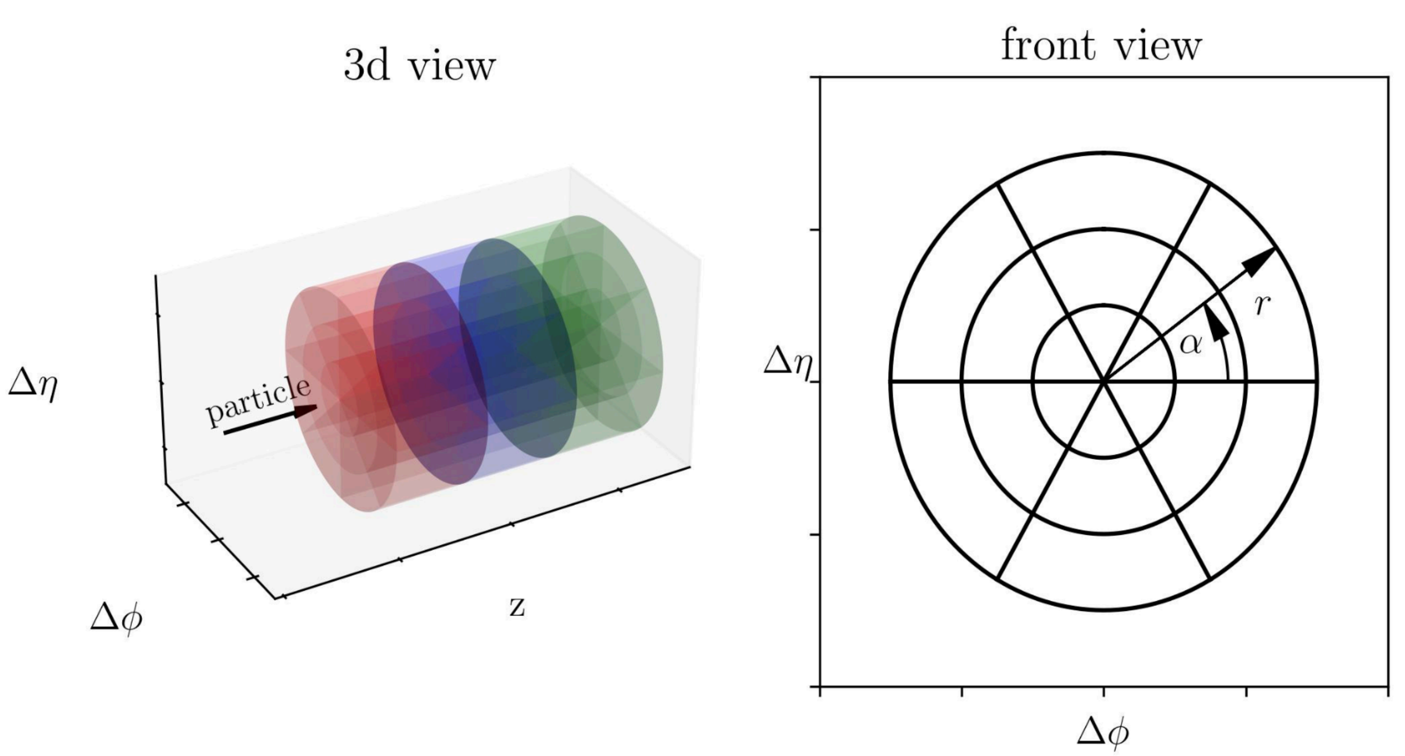
Savannah Thais, Paolo Calafiura, Grigorios Chachamis, Gage DeZoort, Javier Duarte, Sanmay Ganguly, Michael Kagan, Daniel Murnane, Mark S. Neubauer, Kazuhiro Terao

Open data for ML R&D at colliders

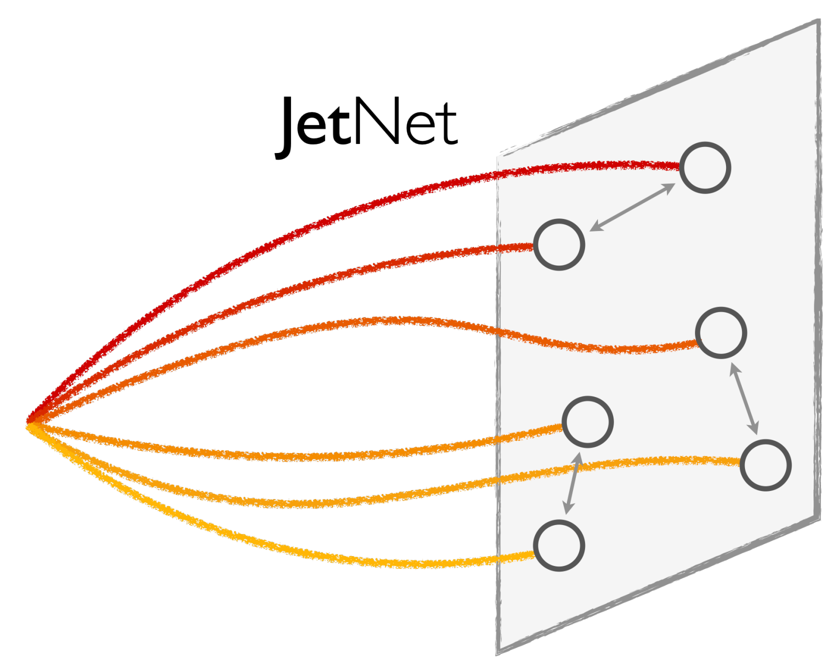


Track-ML challenge :
https://sites.google.com/site/trackmlparticle/home

Calo-ML challenge :
https://sites.google.com/site/trackmlparticle/home



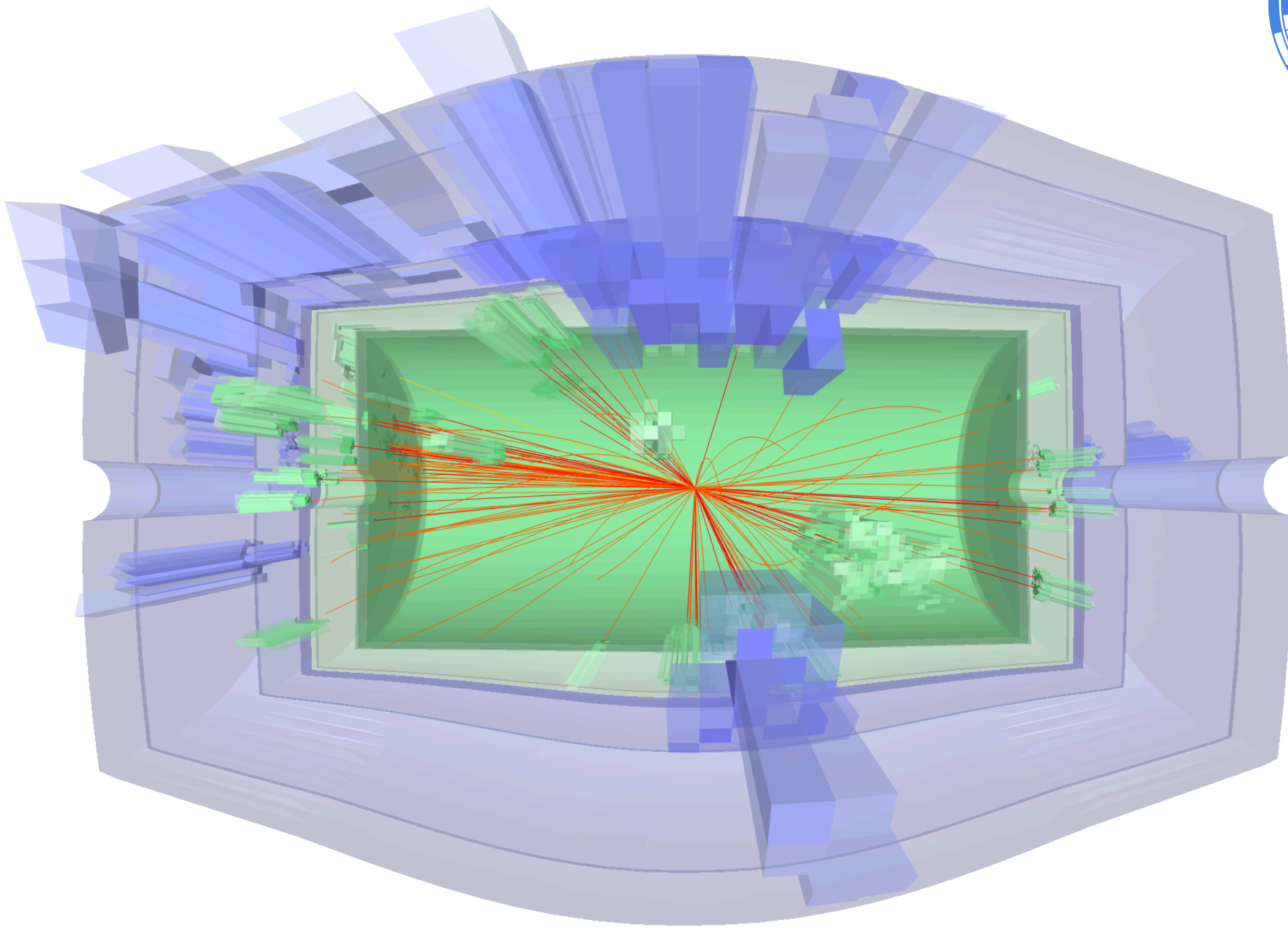
https://github.com/rkansal47/JetNet

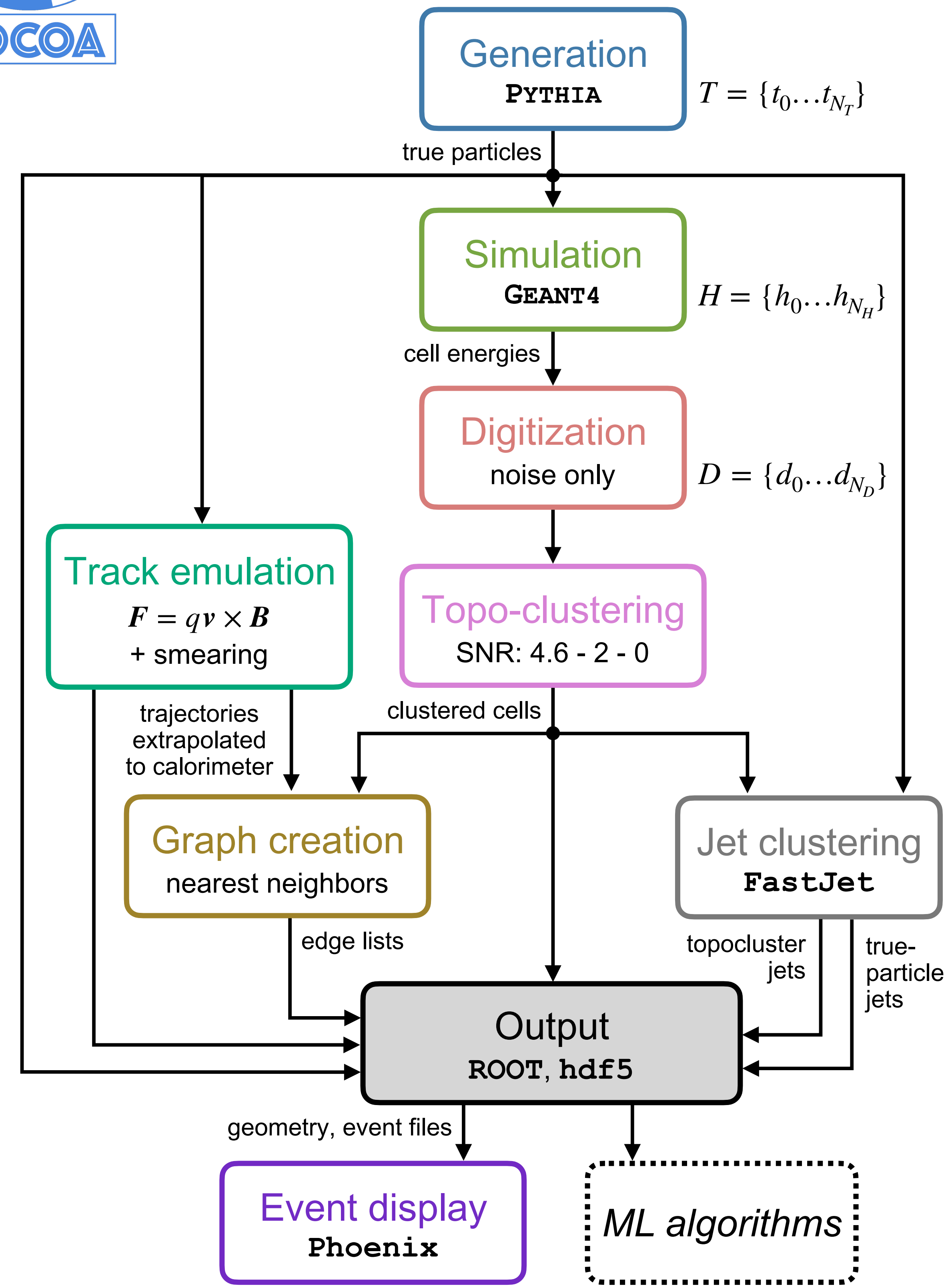


The COCOA

Mach. Learn.: Sci. Technol. 4 035042

https://cocoa-hep.readthedocs.io/en/latest/



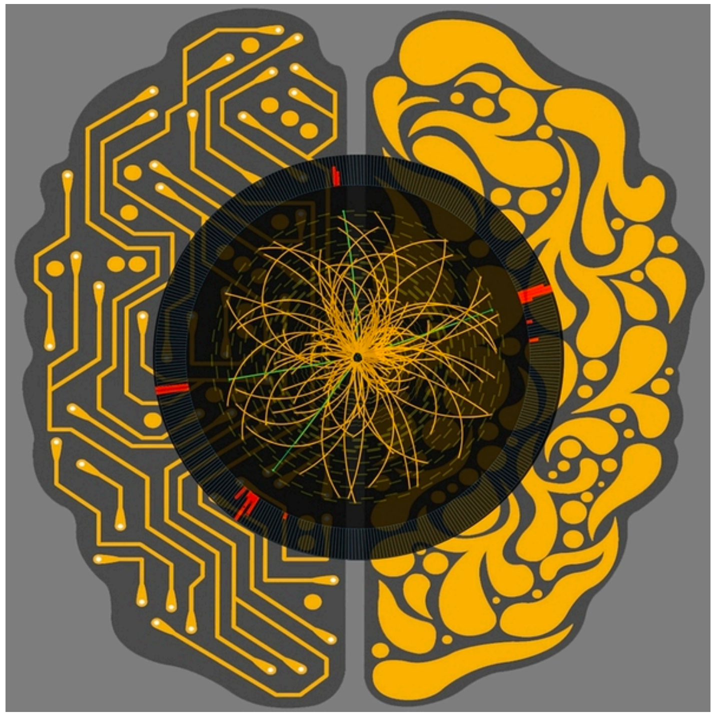


COnfigurable **C**al**O**rimeter simulation for **A**I

- ☑ **A complete hermetic geometry with full GEANT simulation.**
- ☑ **PYTHIA-8 based ME/PS & Hadronization**
- ☑ **FASTJET integration is inbuilt.**
- ☑ **Comes with an ATLAS style pPFlow.**

Take away

Image: FermiLab



- ML is here to stay with HEP.
- We can’t blindly do a plug & play of the available NN.
- Interpretability and uncertainty estimations are two key aspects where we the HEP-ML people need to emphasize.
- Need to keep a close connection with the comp-sc/math community with the latest developments and contribute if possible.
- Symmetry equivariance and geometric DL methods might play a key role in this field.
- Many important application of ML are happening in hep-lat and hep-th community as well.

https://iml-wg.github.io/HEPML-LivingReview/