Graph Neural Network Based Flavor Tagging at LHC

Sanmay Ganguly
(sanmay@icepp.s.u-tokyo.ac.jp)

Advanced jet flavour tagging : ILC

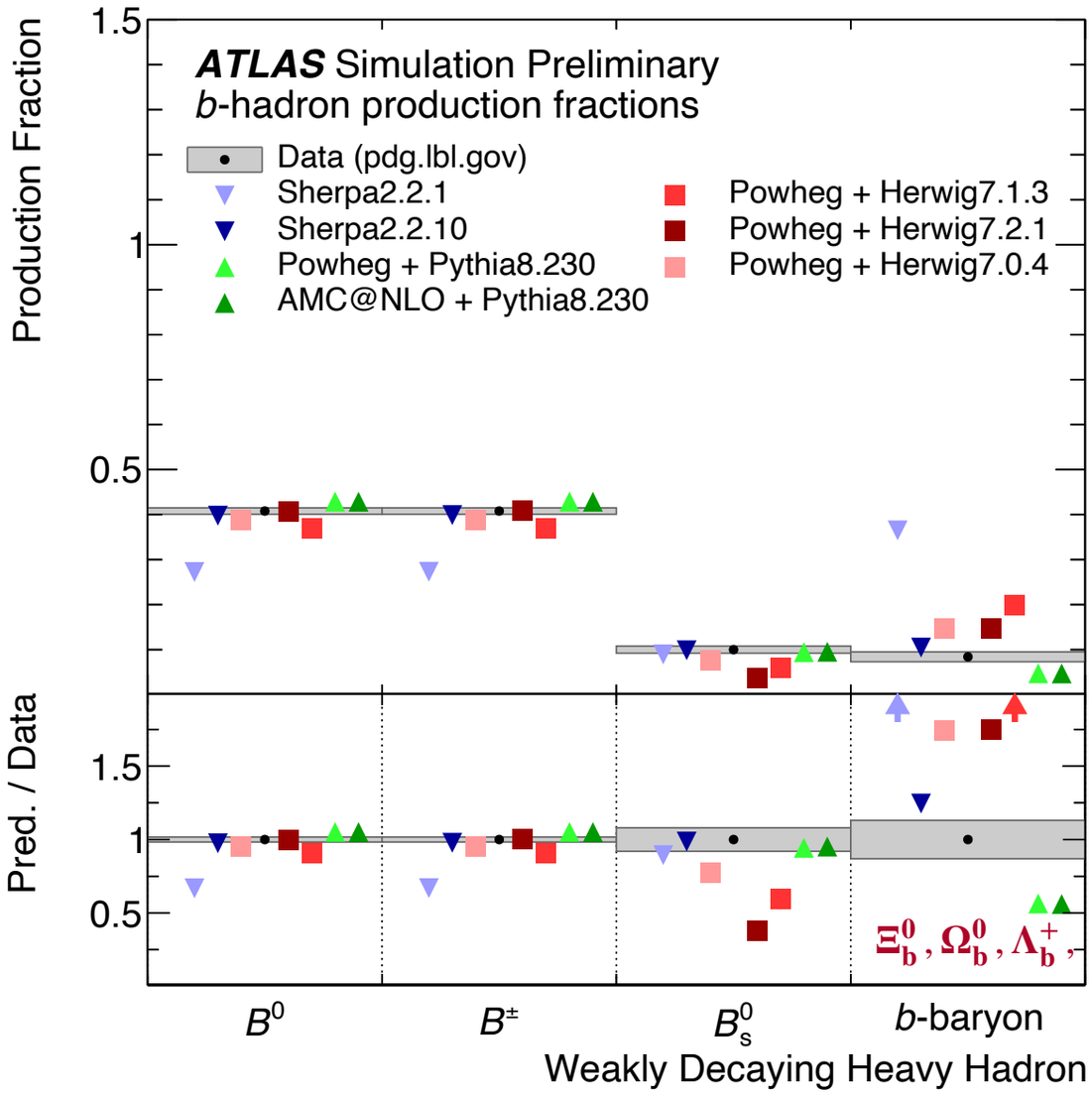
15/03/2023

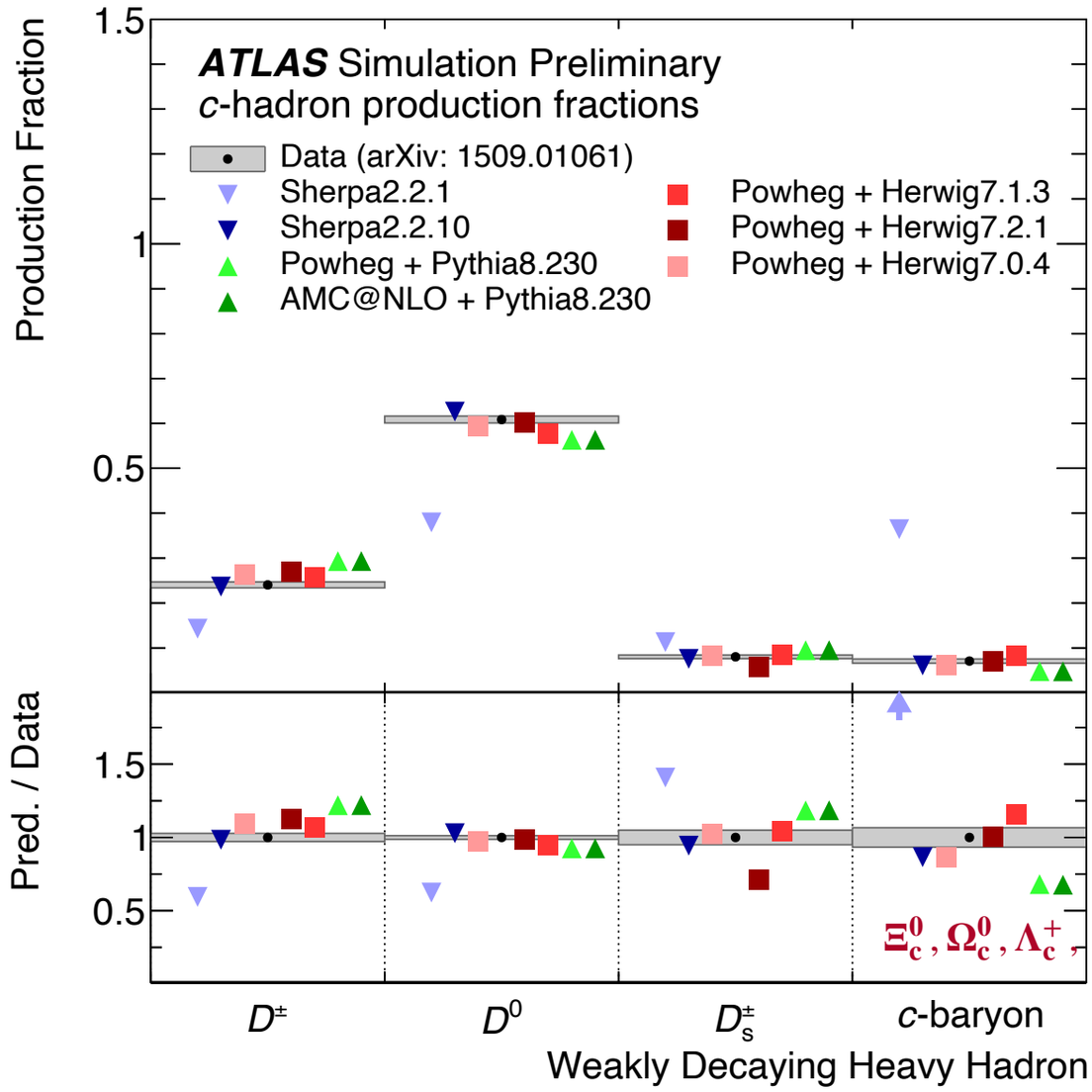
ICEPP The University of Tokyo

Beyond AI

Anatomy of heavy-quark hadronization

How bottom/charm quark decays ?
ATL-PHYS-PUB-2022-035



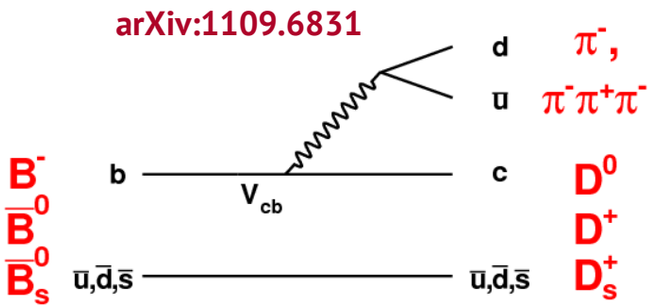


B-hadrons often decays to (D-hadrons + X) via b to c transition.
D-hadrons further decays to bunch of Kaons, Pions, leptons etc.

Anatomy of heavy-quark hadronization

The causal direction

arXiv:1109.6831



$\beta\ \gamma\ \tau_B \sim 1/|V_{bc}|^2$

$\langle \tau_B \rangle \sim$ 1.566 *ps*

K^+

l

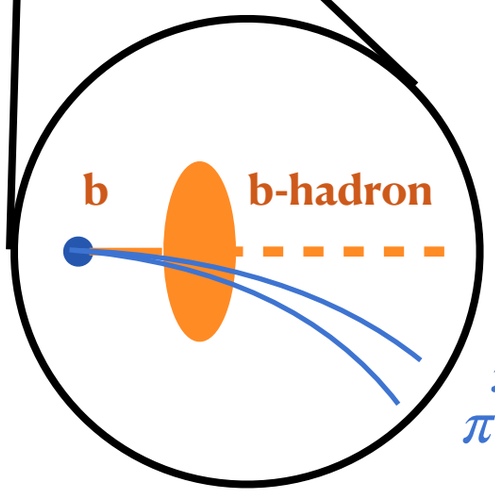
π^-

$\langle \tau_D \rangle \sim$ 0.1 – 1 *ps*

l

K^+

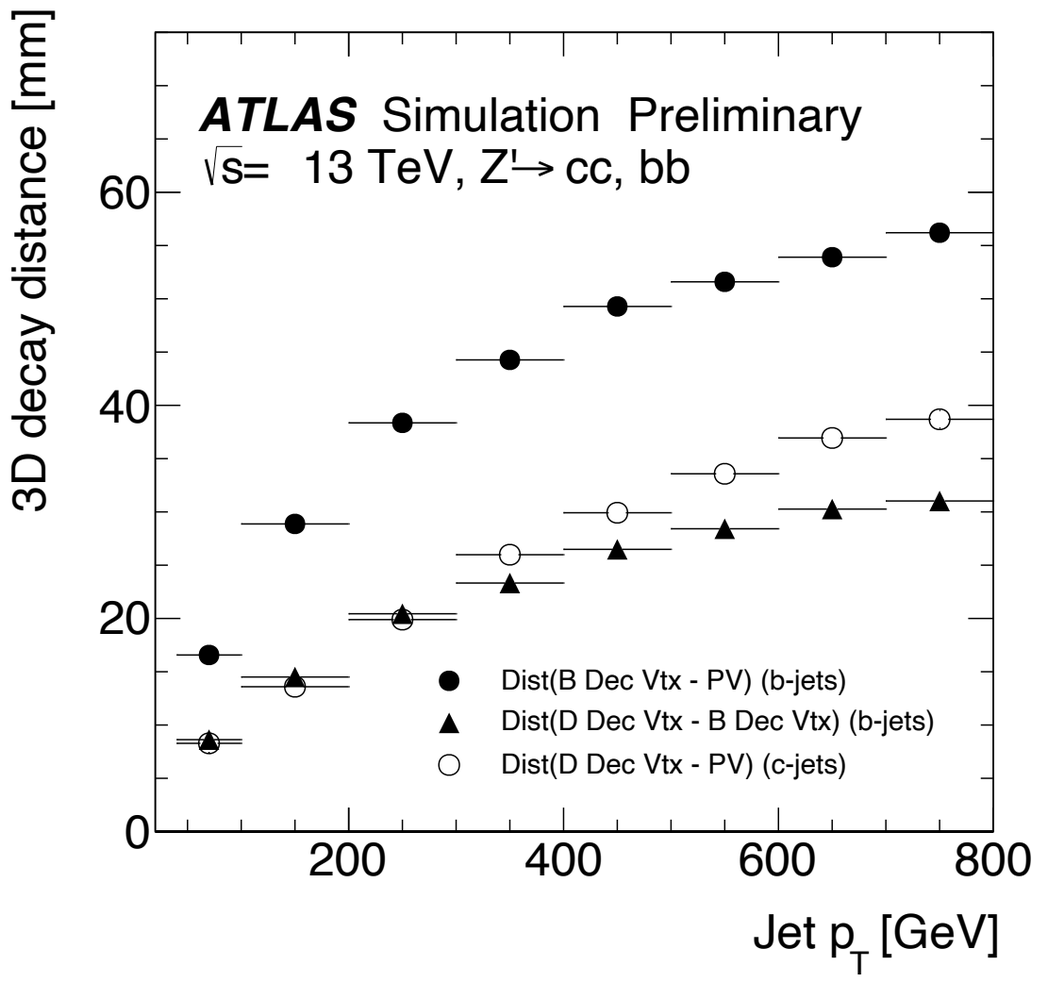
π^-



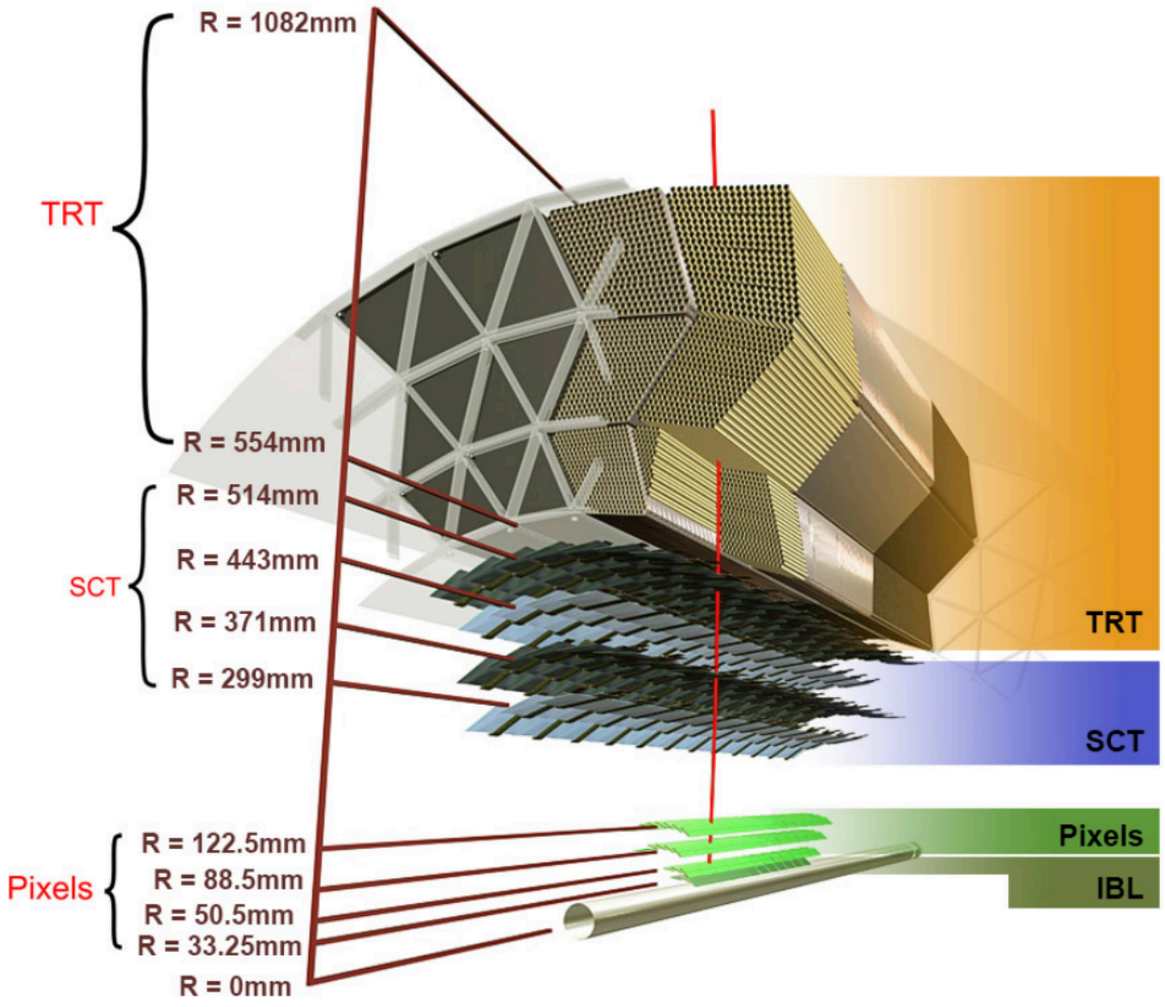
$\tau_{had} \sim \Lambda_{QCD}^{-1} \sim 10^{-24}$ sec

The decay topology in detector

ATL-PHYS-PUB-2018-025



arXiv : 2007.07624

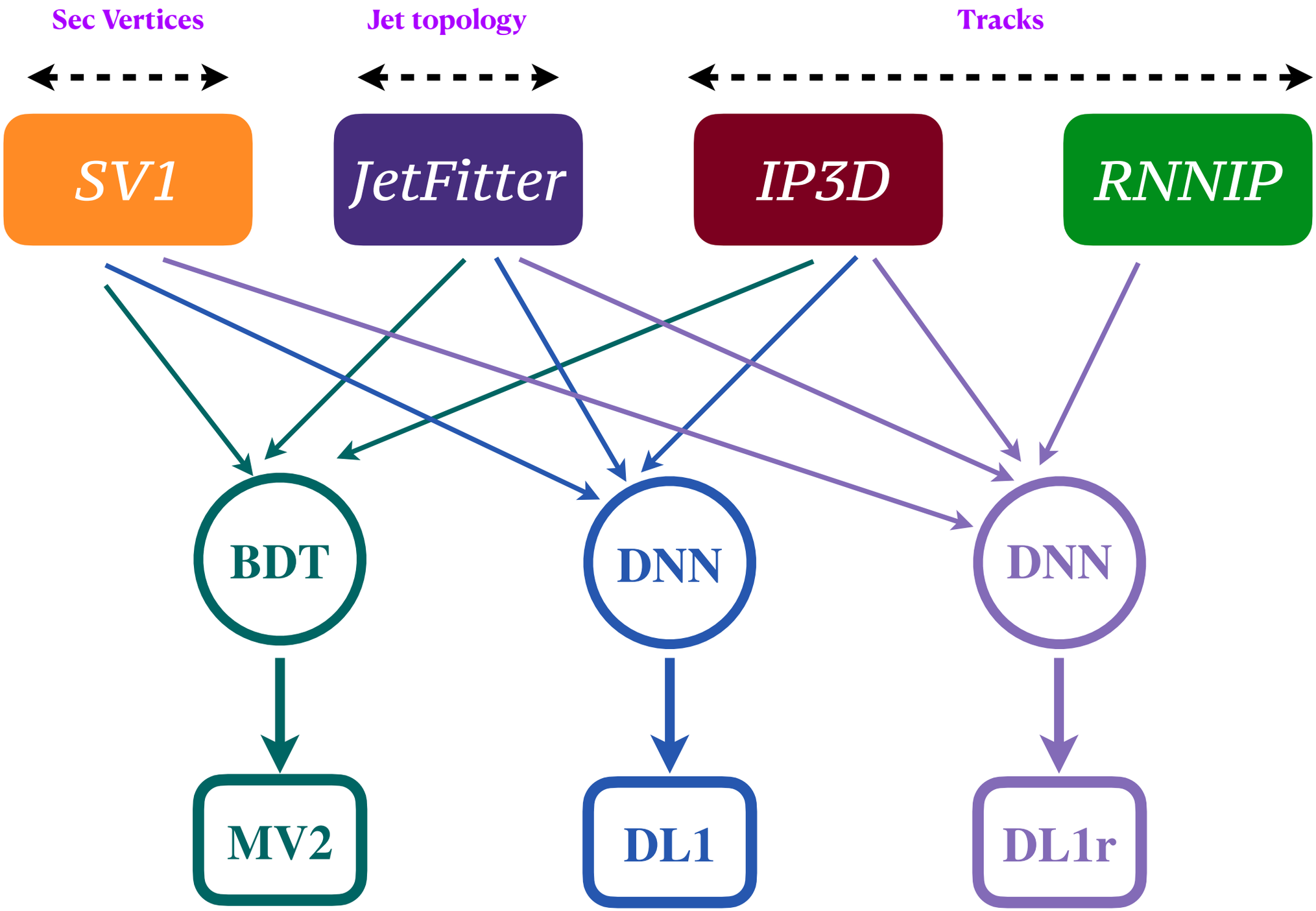


Track : constructed from ID and pT > 500 MeV.

Primary vertex : every event should have one with $\sigma(z) \sim 30\ \mu m$ and $\sigma(r) \sim 10\ \mu m$

Jets : reconstructed with Ak4 algorithm, with $p_T > 20$ GeV and $|\eta| < 4.5$ + a jet-PV association.

High level algorithms

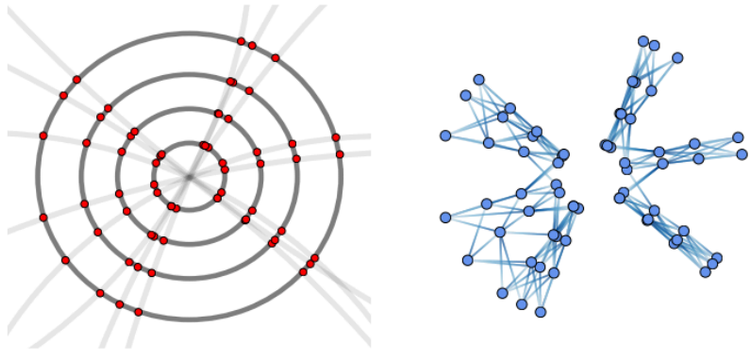


MV2 vs DL1 : same input, different architecture.

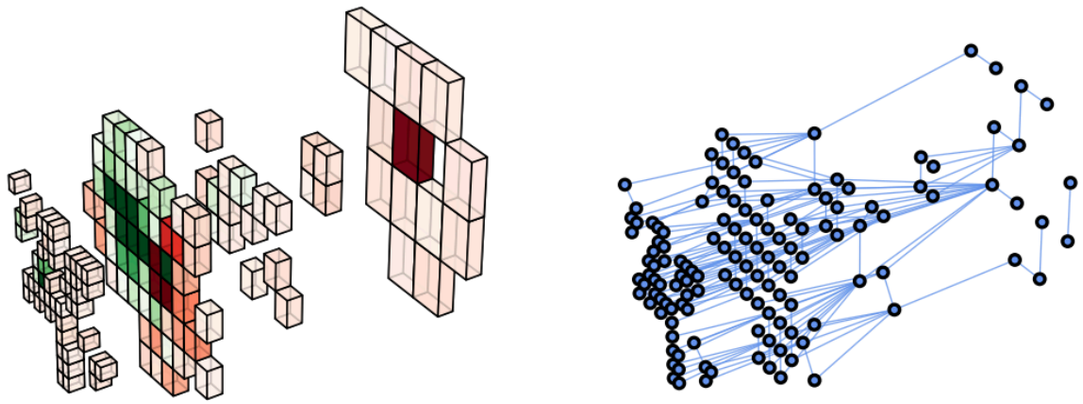
DL1r has extra RNNIP input compared to DL1.

HEP events & objects as graph

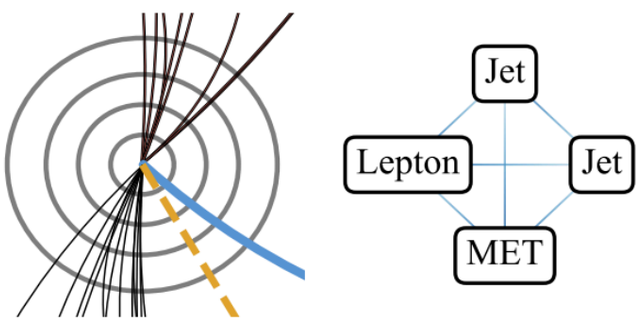
arXiv : 2007.13681



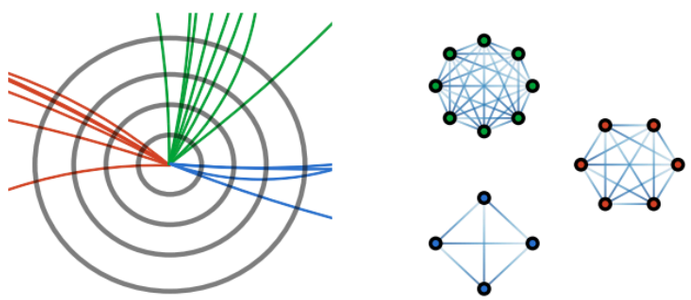
(a)



(b)

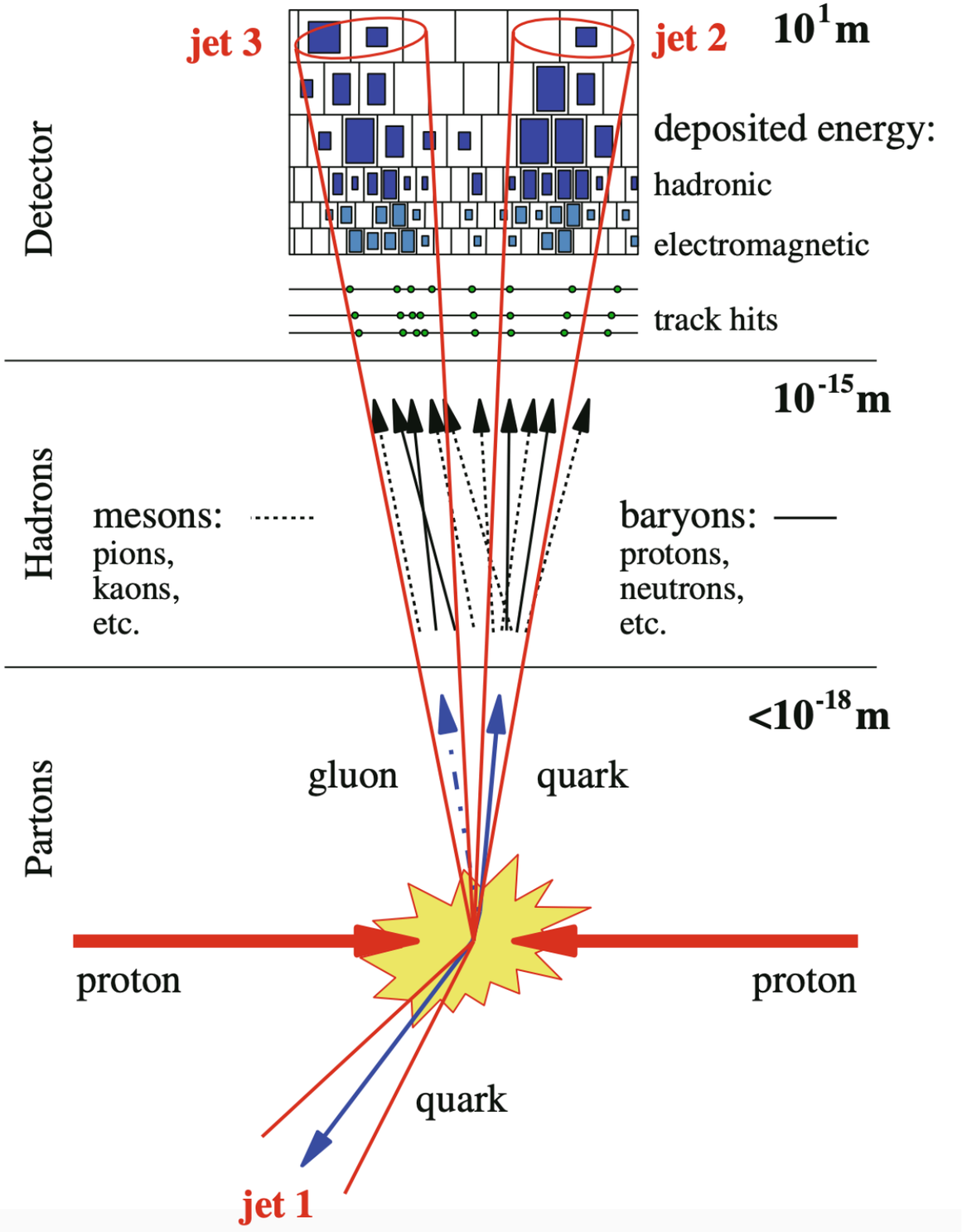


(c)



(d)

Jet tagging ~ Graph classification



The Large Hadron Collider
Editors: Thomas Schörner-Sadenius

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

Jet Algorithm (for CA, kT, anti-kT)

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

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

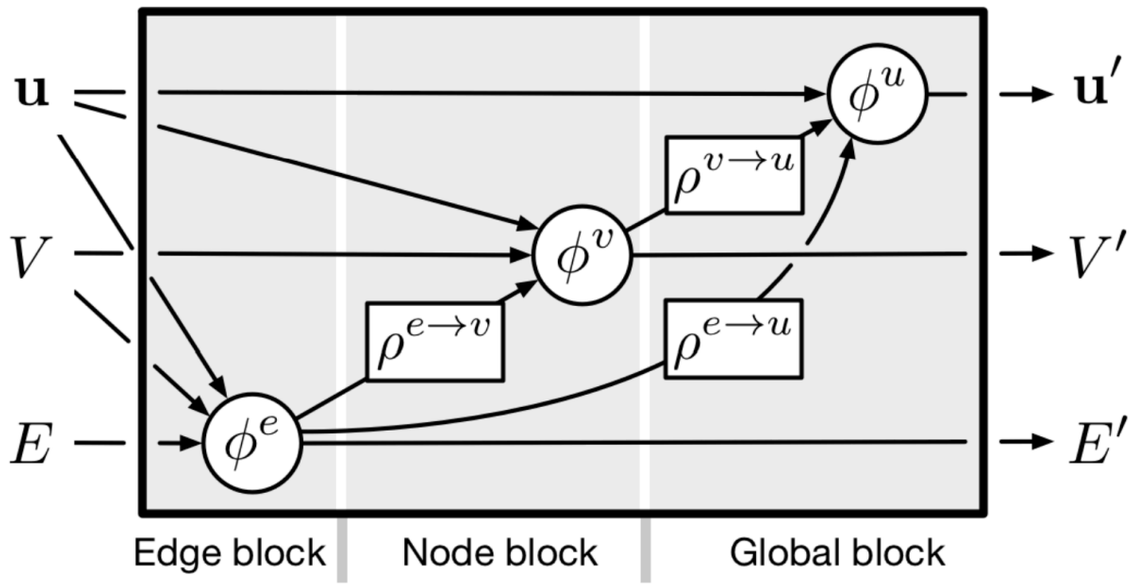
The forward problem is not computable from first principle

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

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

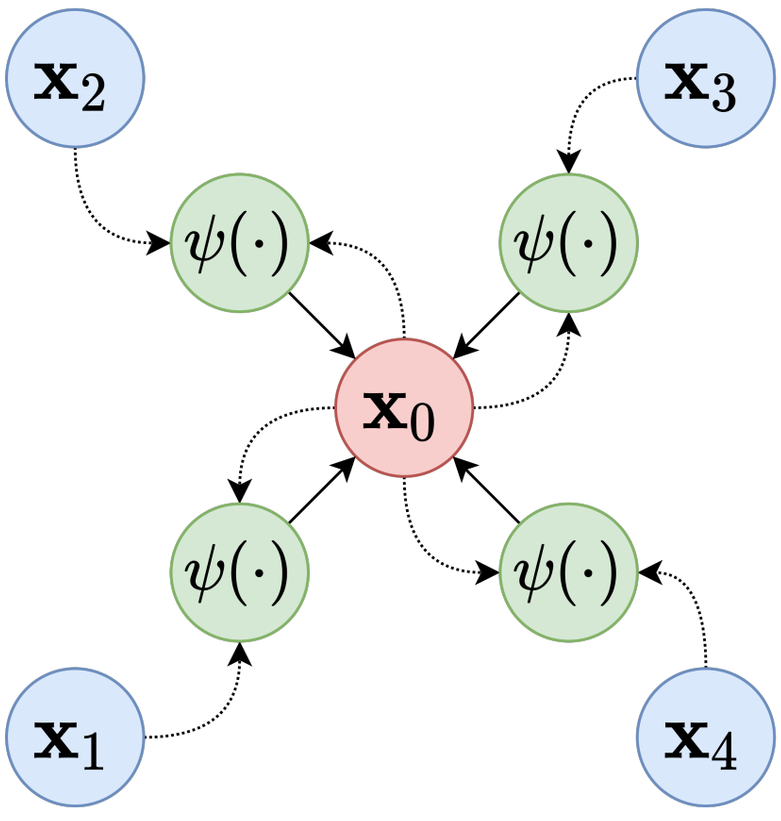
What is GNN?

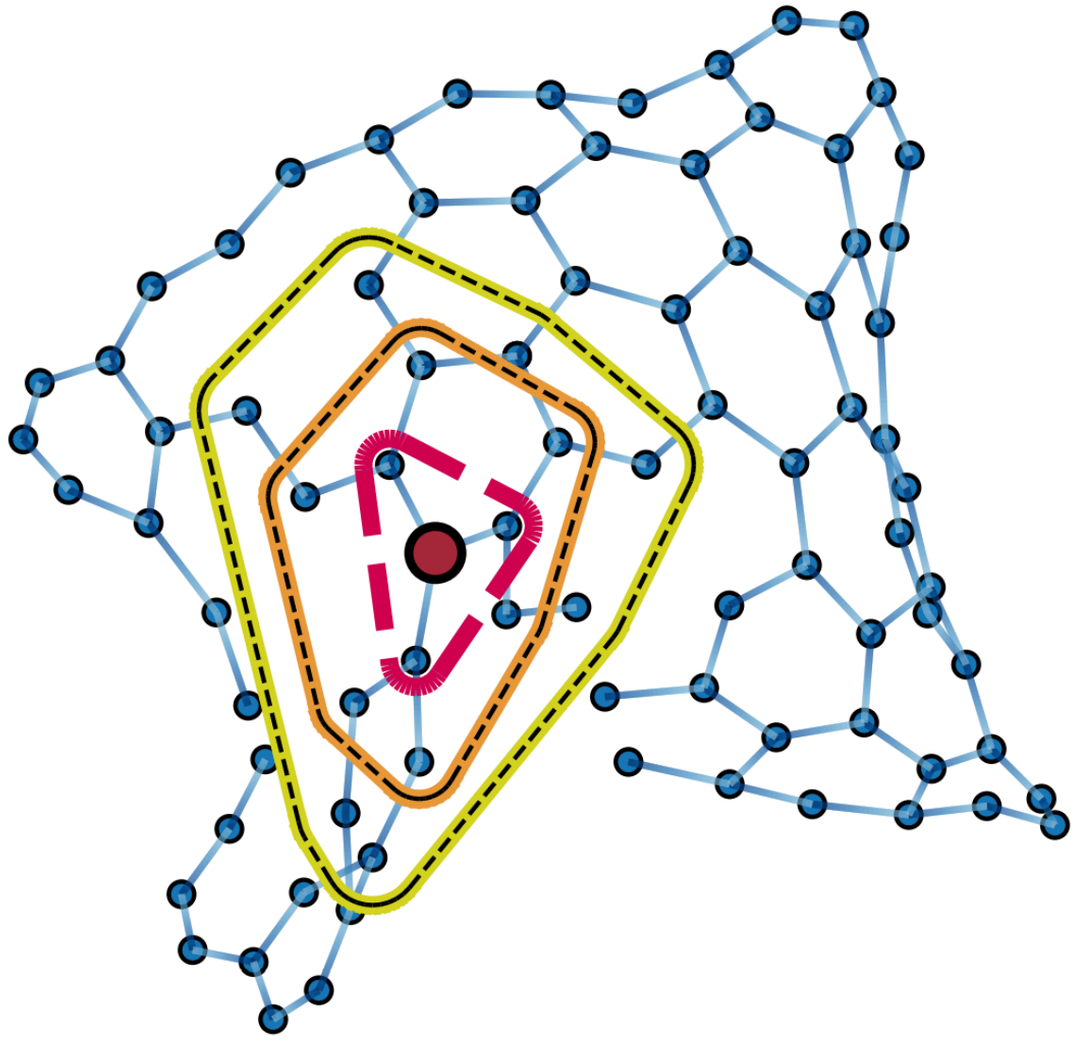
$$G = (u, V, E)$$



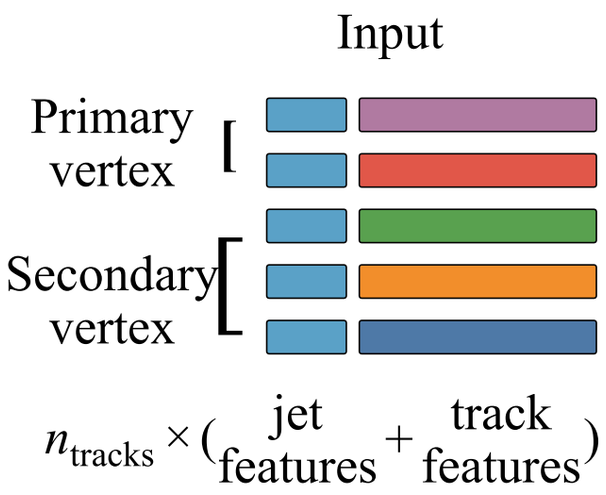
arXiv : 1806.01261

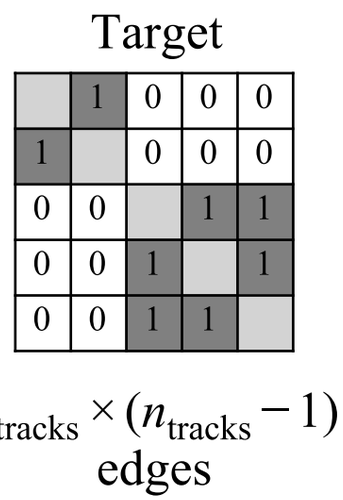
$$x_i' = \phi\left(x_i, \oplus_{i\in\mathcal{N}_i}\psi(x_i, x_j, e_{ij})\right)$$

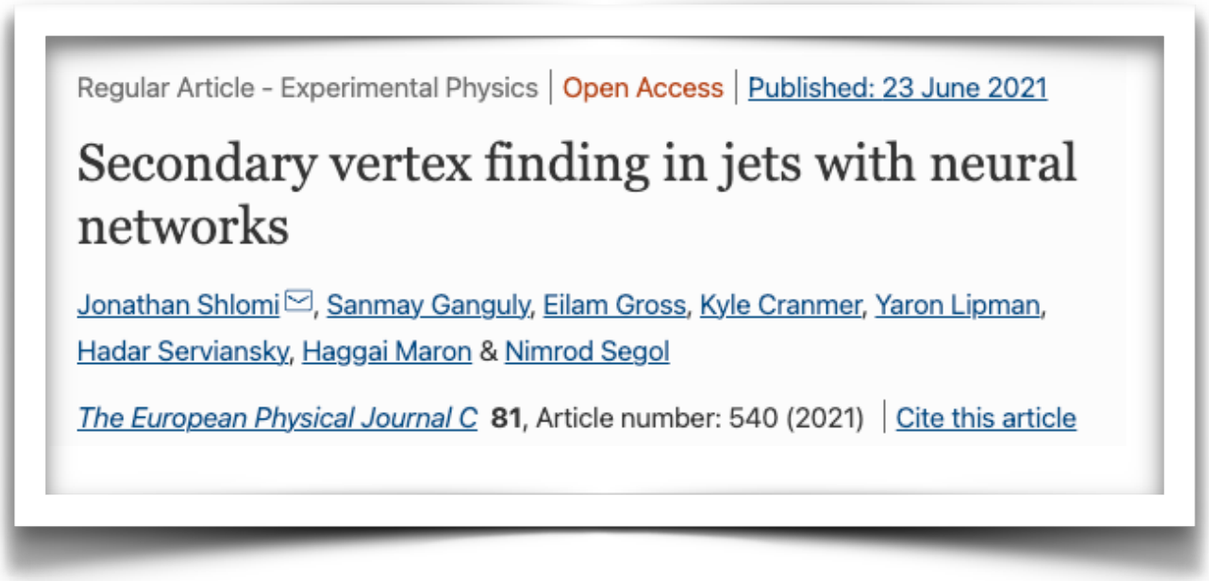


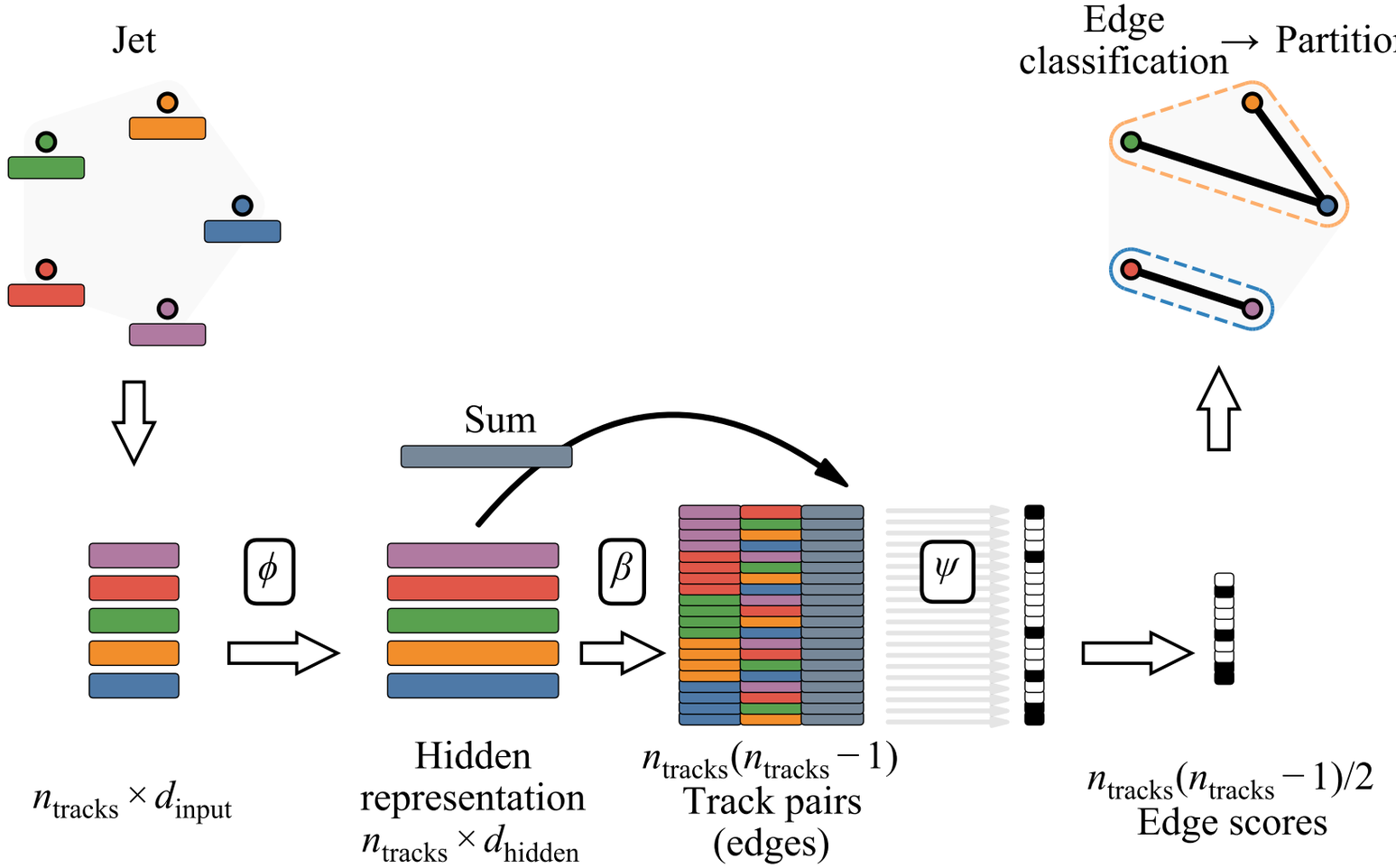


Set2Graph proposal for flavor-tagging

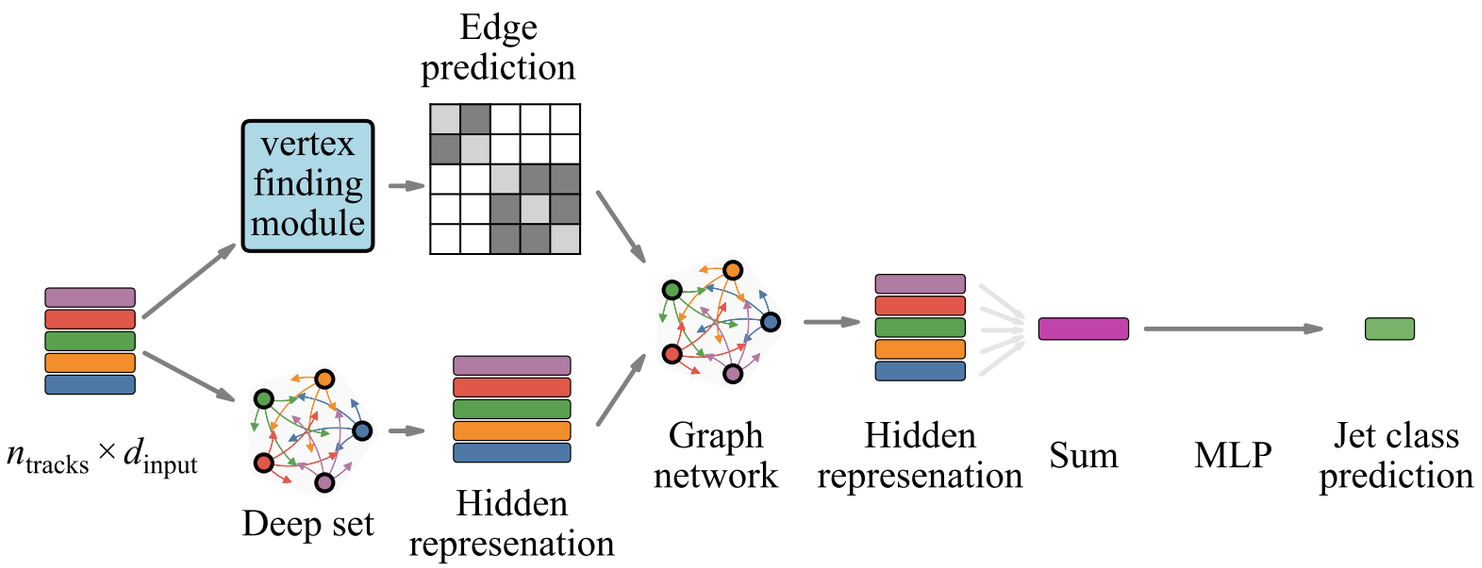


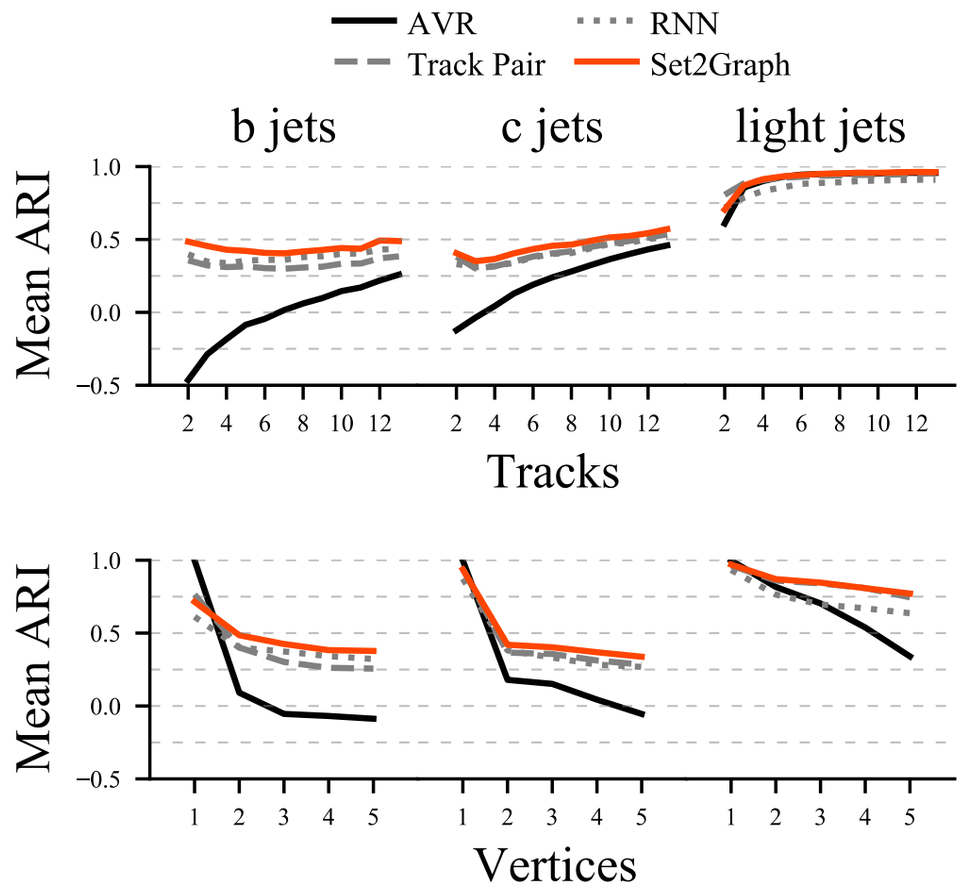






Set2Graph proposal for flavor-tagging





$$\text{RI} = \frac{\text{number of correct edges}}{\text{number of edges in the set}}$$

$$\mathbb{E}[\text{RI}] = \frac{B_{N-1}}{B_N}\frac{\sum_i \binom{g_i}{2}}{\binom{N}{2}} + \left(1 - \frac{B_{N-1}}{B_N}\right)\left(1 - \frac{\sum_i \binom{g_i}{2}}{\binom{N}{2}}\right)$$

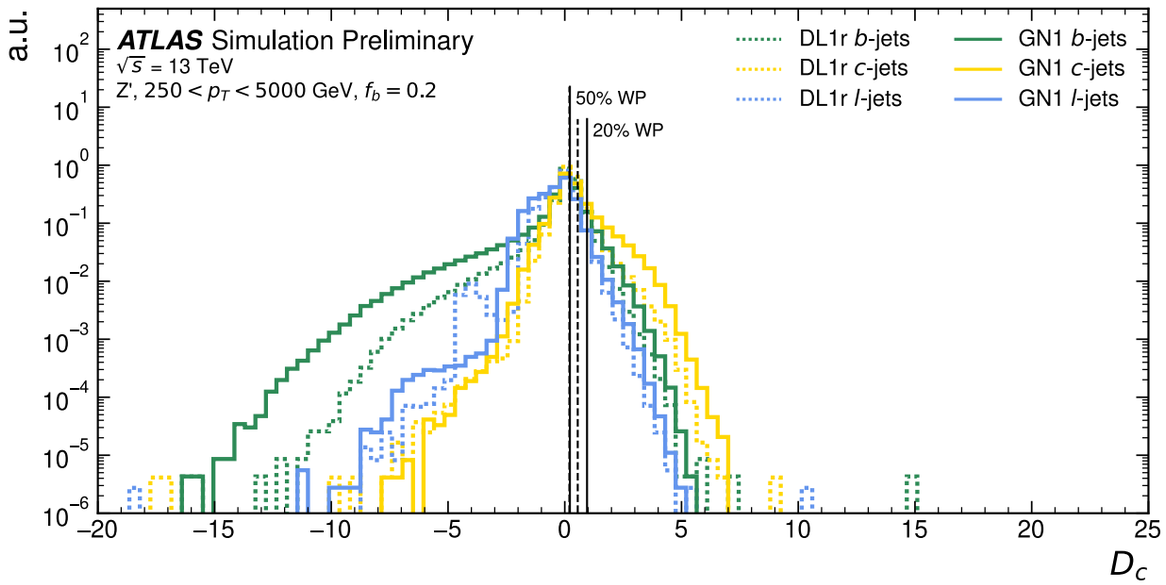
$$\text{ARI} = \frac{\text{RI} - \mathbb{E}[\text{RI}]}{1 - \mathbb{E}[\text{RI}]}$$

	Algorithm	Jet			Vertex			Vertex-Pair				
		F1	RI	ARI	internal	external	combined	internal_1	internal_2	interpair	external	combined
b jets	AVR	0.56	0.61	-0.01	**0.91**	0.51	0.46	**0.59**	**0.90**	0.54	0.58	0.18
	Track Pair	0.62	0.74	0.32	0.86	0.71	0.60	0.55	0.87	0.72	0.74	0.29
	RNN	0.59	0.75	0.37	0.79	**0.77**	0.60	0.48	0.84	**0.78**	**0.80**	0.27
	Set2Graph	**0.66**	**0.78**	**0.43**	0.86	0.76	**0.65**	0.54	0.88	**0.78**	0.79	**0.33**
c jets	AVR	0.70	0.65	0.22	**0.95**	0.41	0.39	**0.49**	**0.91**	0.49	0.66	0.14
	Track Pair	0.74	0.73	0.40	0.92	0.58	0.52	0.47	0.88	0.65	0.76	0.24
	RNN	0.71	0.72	0.40	0.86	**0.60**	0.50	0.39	0.85	0.65	0.77	0.19
	Set2Graph	**0.75**	**0.75**	**0.45**	0.94	**0.60**	**0.56**	0.47	**0.91**	**0.67**	**0.78**	**0.26**
light jets	AVR	**0.97**	**0.96**	0.93	**0.99**	0.89	0.88	**0.33**	**0.98**	0.73	0.89	0.14
	Track Pair	0.96	**0.96**	0.93	0.97	**0.93**	0.90	0.32	0.97	0.87	**0.95**	**0.26**
	RNN	0.93	0.92	0.87	0.93	0.90	0.84	0.25	0.94	0.82	0.93	0.18
	Set2Graph	**0.97**	**0.96**	**0.94**	0.98	**0.93**	**0.91**	0.32	**0.98**	**0.88**	**0.95**	**0.26**

B_N : # of possible partition of a set with N elements

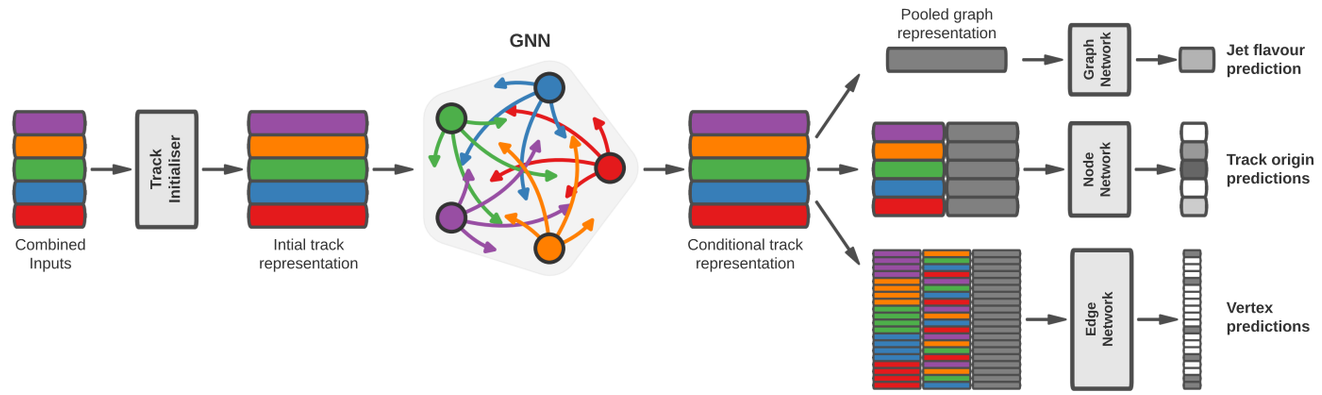
g_i : # of tracks associated with i-th vertex in the jet.

Set2Graph model within ATLAS



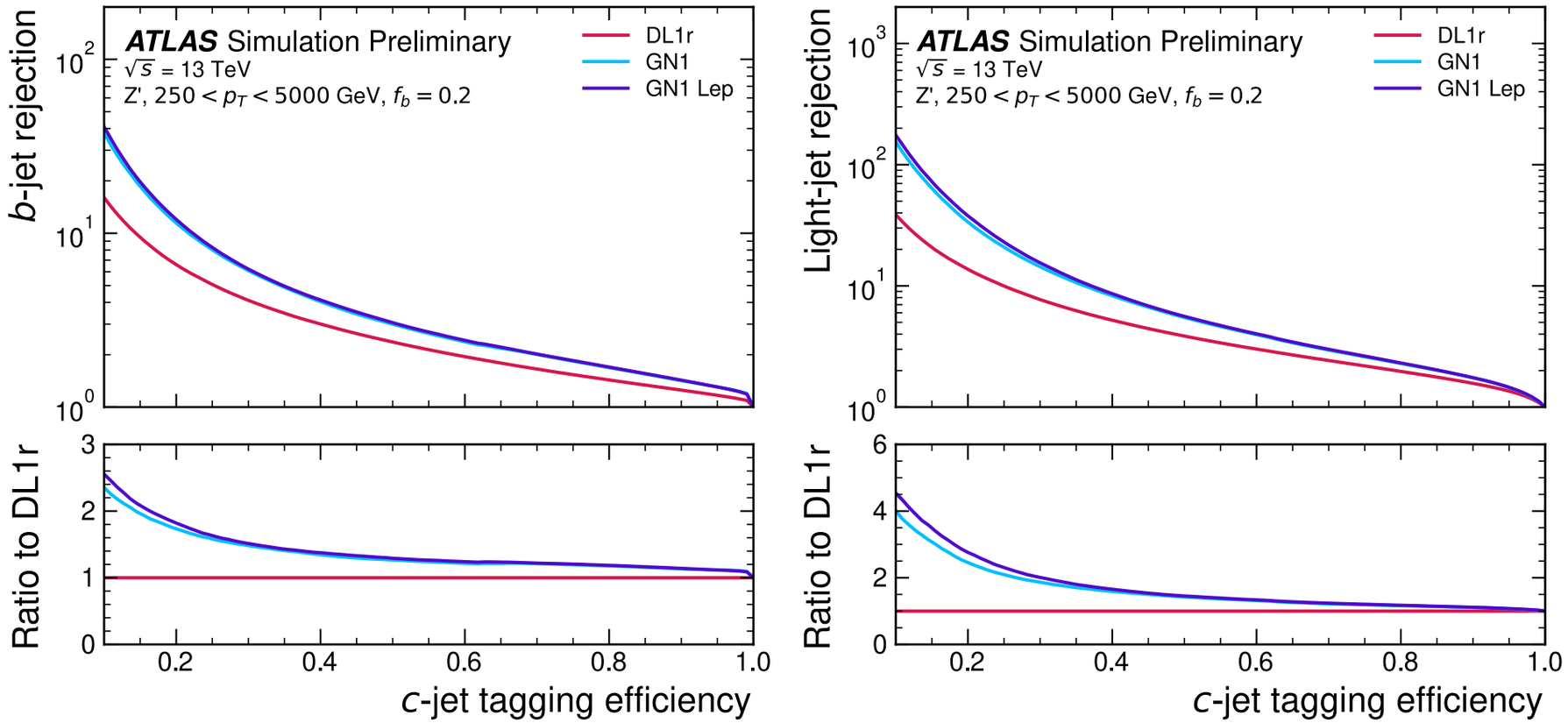
$$D_{DL1}^{b-tag} = ln\left(\frac{p_b}{f_c \cdot p_c + (1-f_c) \cdot p_l}\right)$$

ATL-PHYS-PUB-2022-027



Sizable improvement over the current DL1r algorithm.

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

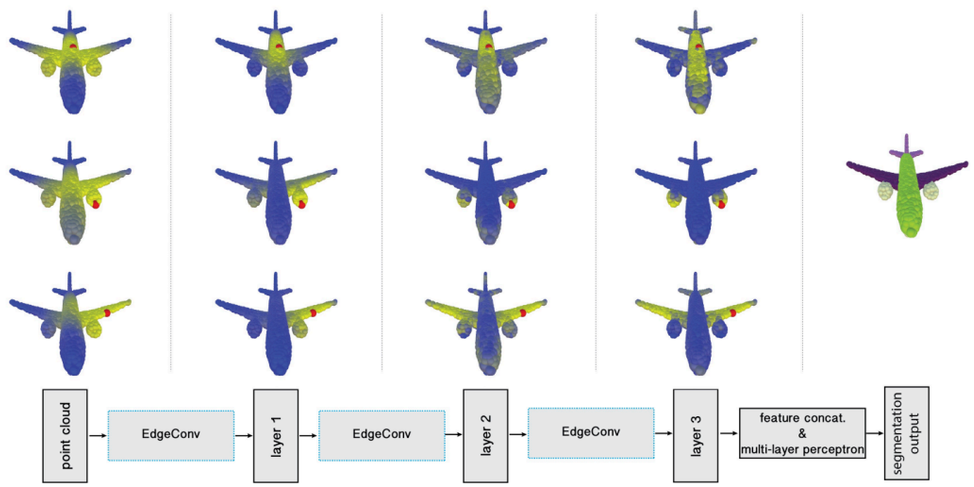


The pointcloud classification

Dynamic Graph CNN for Learning on Point Clouds

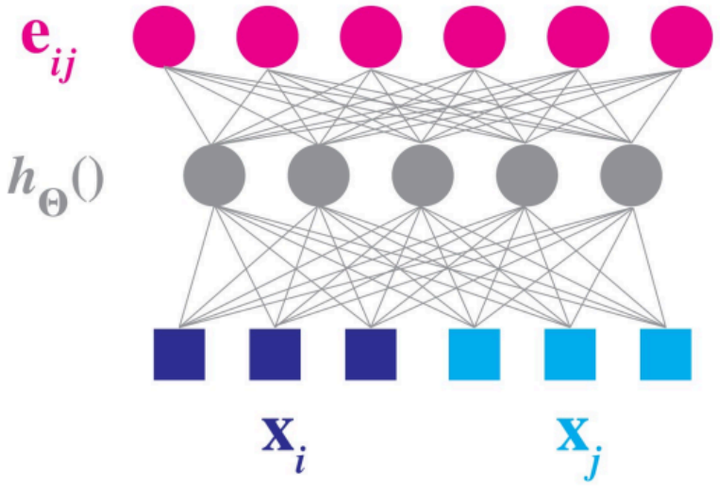
YUE WANG, Massachusetts Institute of Technology
YONGBIN SUN, Massachusetts Institute of Technology
ZIWEI LIU, UC Berkeley / ICSI
SANJAY E. SARMA, Massachusetts Institute of Technology
MICHAEL M. BRONSTEIN, Imperial College London / USI Lugano
JUSTIN M. SOLOMON, Massachusetts Institute of Technology

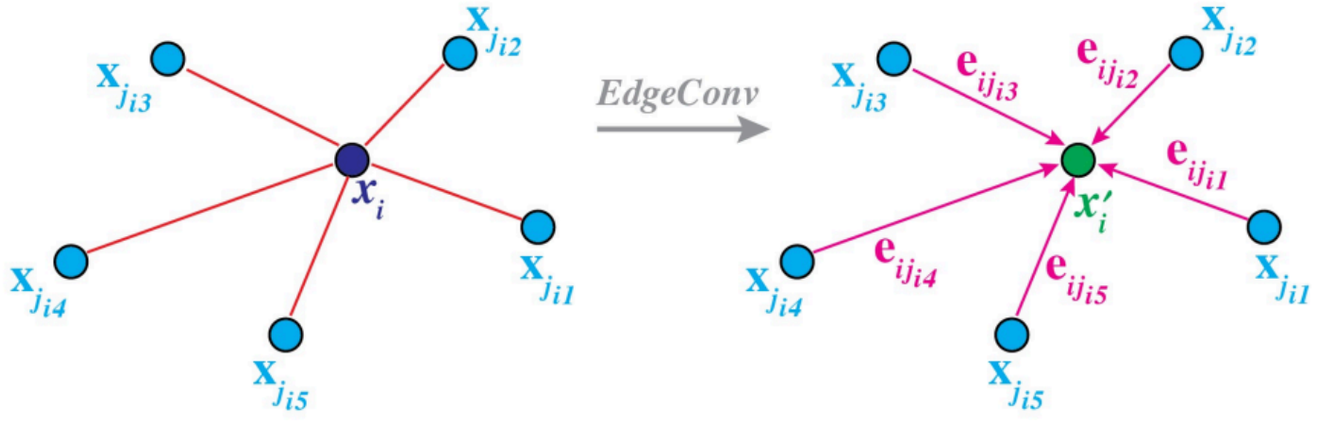
arXiv : 1806.01261



$$\mathbf{x}'_i = \max_{j:(i,j)\in\mathcal{E}} h_{\Theta}(\mathbf{x}_i, \mathbf{x}_j)$$

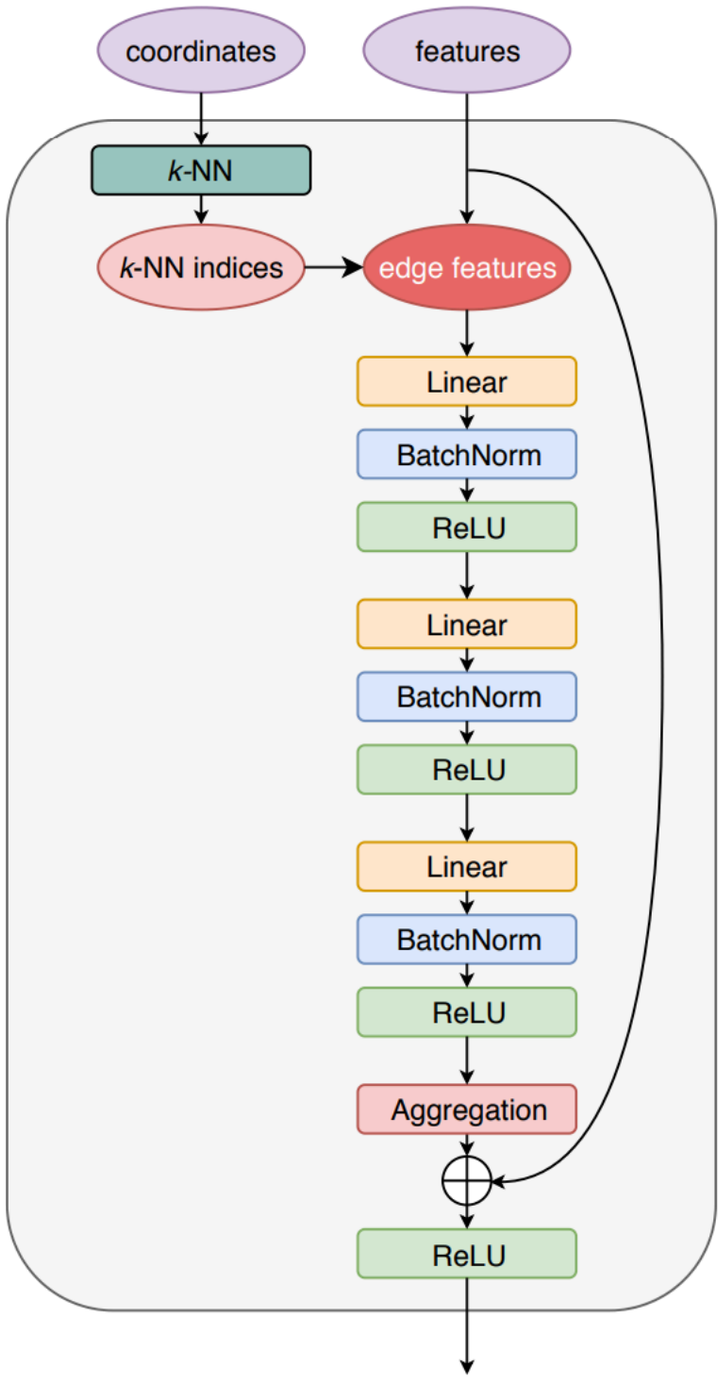
$$e'_{ijm} = \mathrm{ReLU}(\boldsymbol{\theta}_m \cdot (\mathbf{x}_j - \mathbf{x}_i) + \boldsymbol{\phi}_m \cdot \mathbf{x}_i),$$

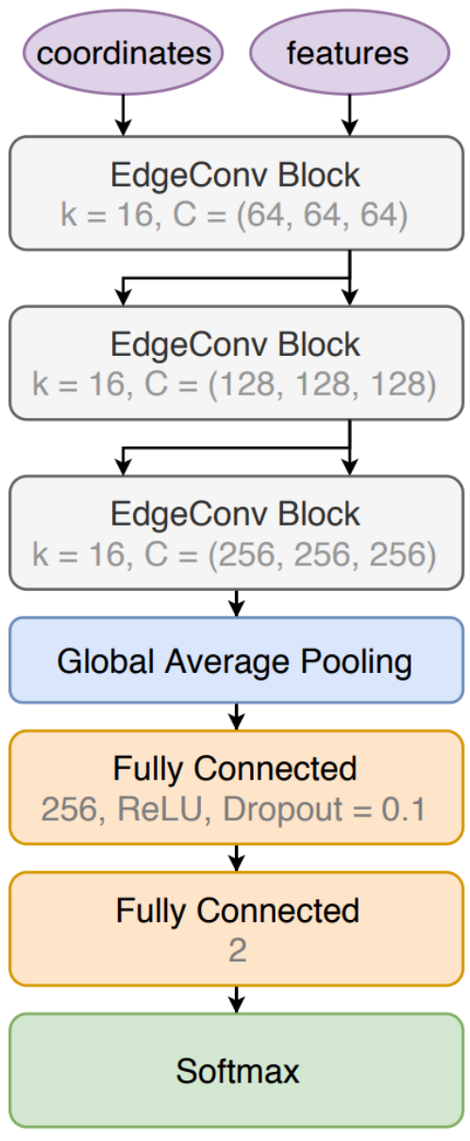




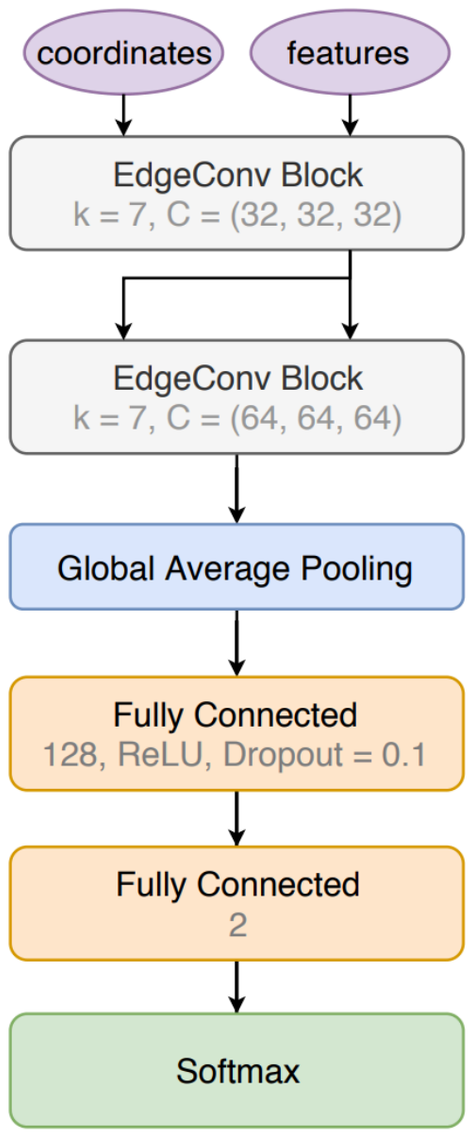
The particlenet network

arXiv : 1902.08570





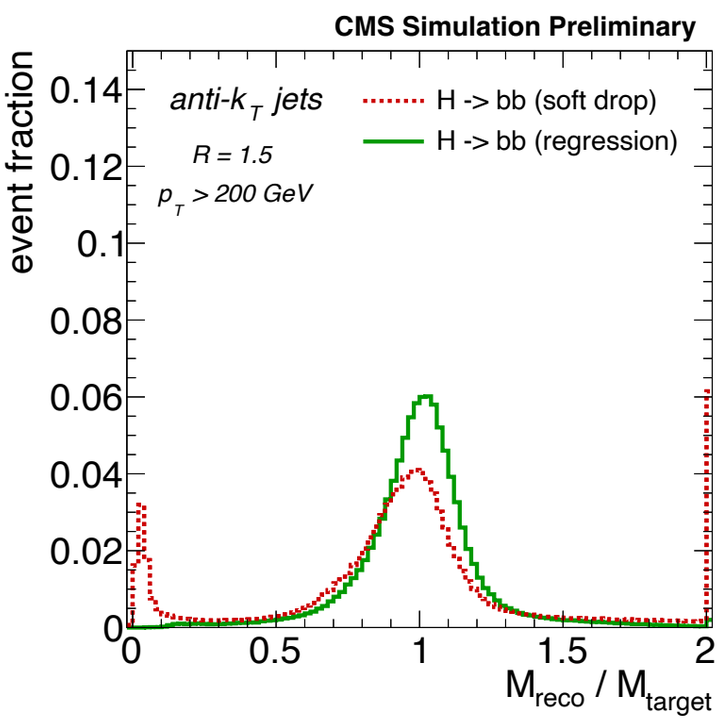
(a) ParticleNet

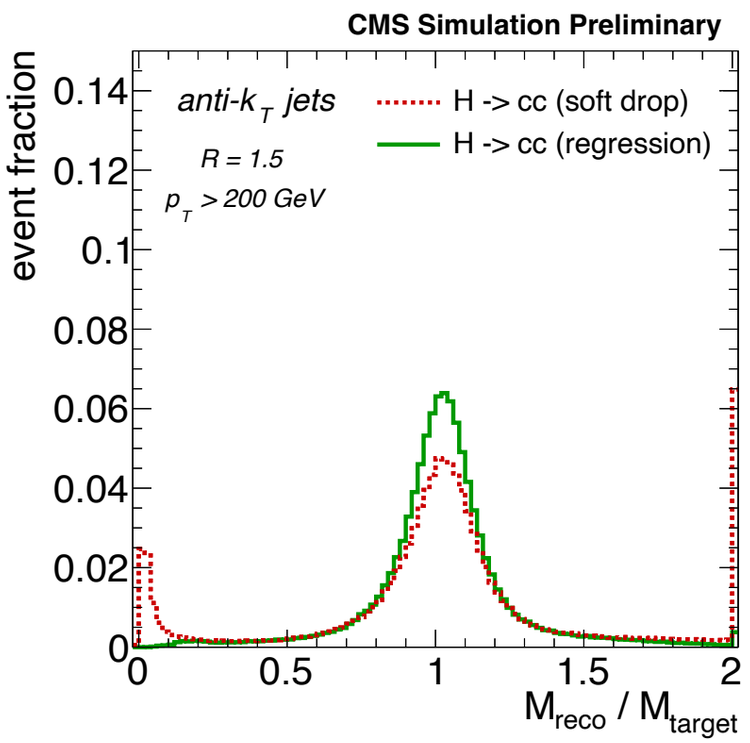


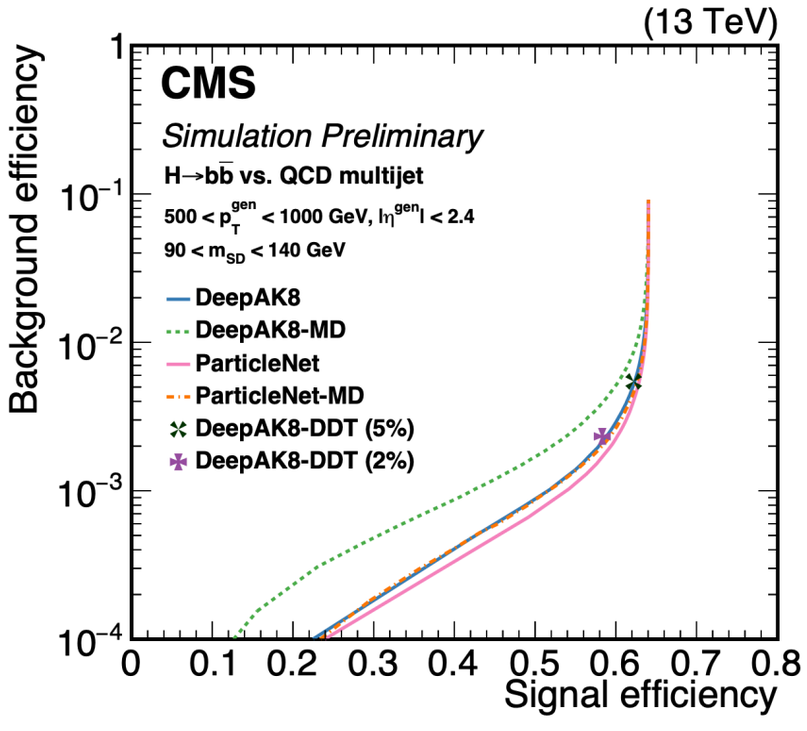
(b) ParticleNet-Lite

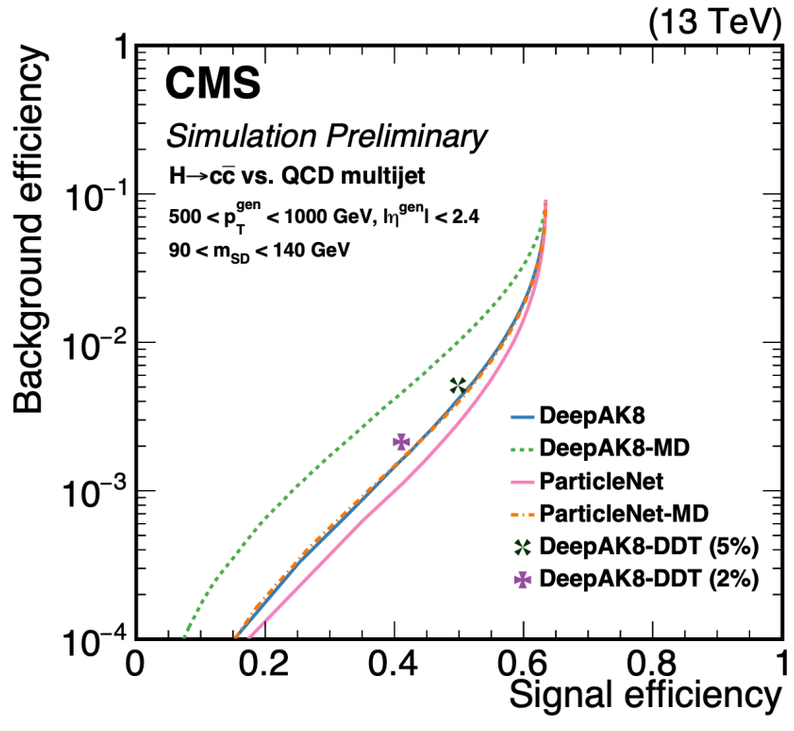
The large-R mass regression

CMS-DP-21017



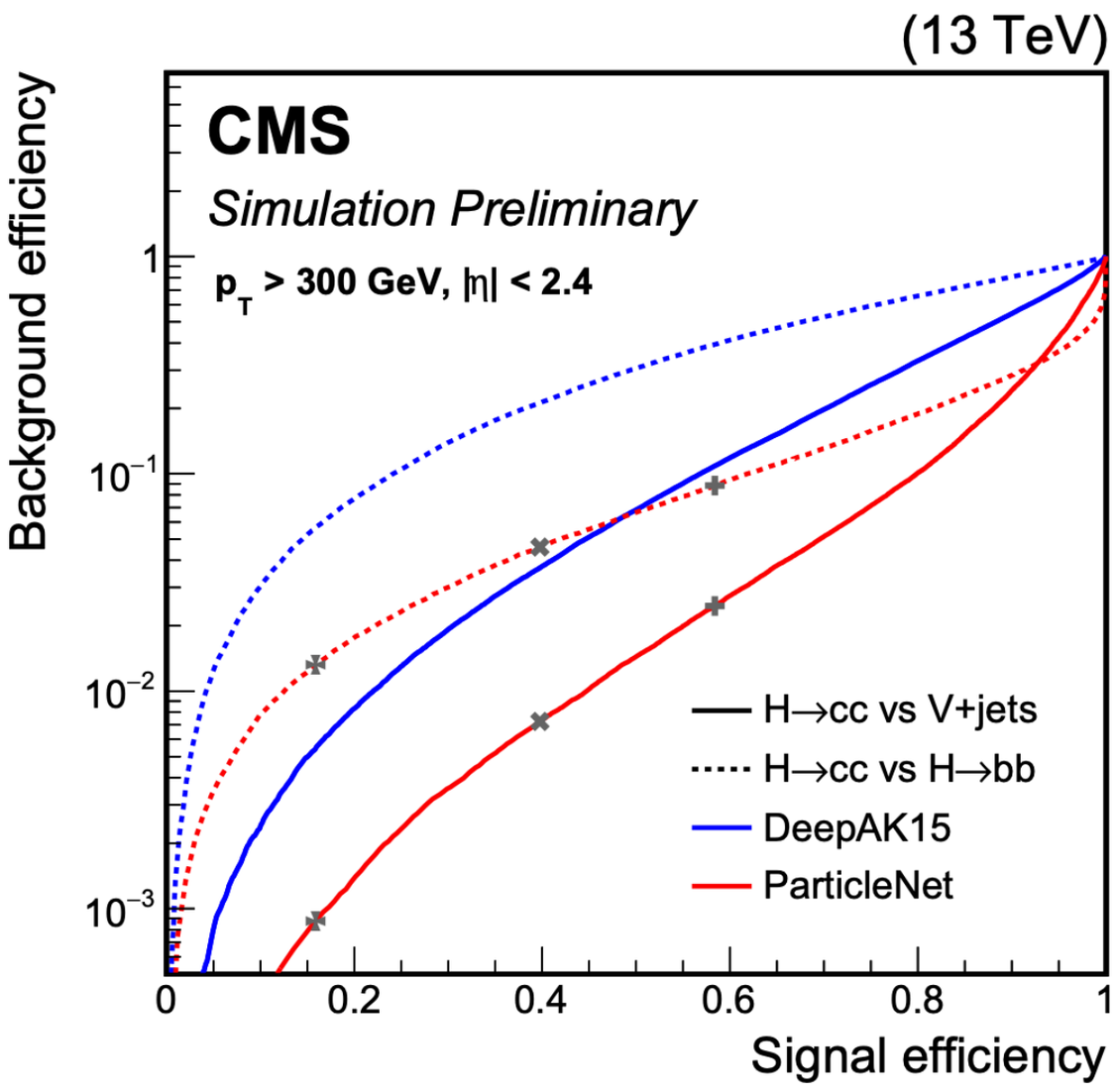


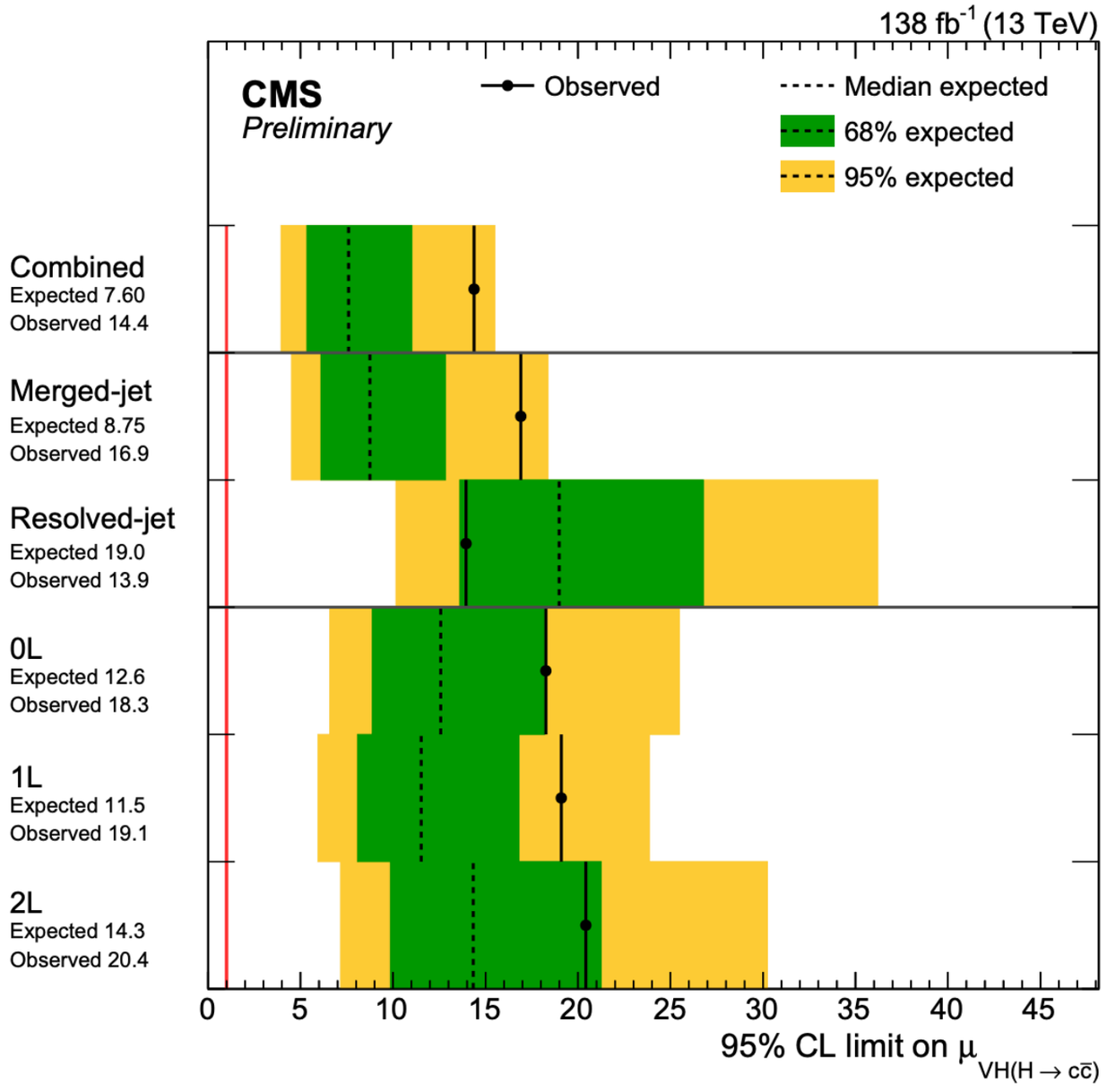




CMS-DP-2020/002

Impact of GNN on physics results





CMS-HIG-21-008

Take away

- ☑ **GNN is turning out to be the state of art ML technique to be used for HEP applications.**
- ☑ **GNN has its own limitation, need to understand how HEP analysis is effected.**
- ☑ **We need to emphasize on explainable AI and correlate to physics interpretation.**
- ☑ **Let's see what it can bring for ILC.**

THANK YOU!!