Boost-Invariant Polynomials:

an efficient and interpretable approach to jet tagging

ACAT 2022

24/10/2022

Personal Introduction

Christoph Ortner

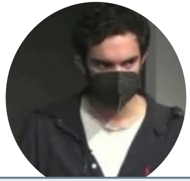
Ilyes Batatia

Jose M Munoz

Personal Introduction

Christoph Ortner

THE UNIVERSITY OF BRITISH COLUMBIA



Ilyes Batatia

UNIVERSITY OF CAMBRIDGE

- Undergrad Physics Student
- ML Engineer at Guane
- Focus on Geometric-DL and ML for HEP
- UBC Intern
- CERN Summer Student

Jose M Munoz

UNIVERSIDAD EIA

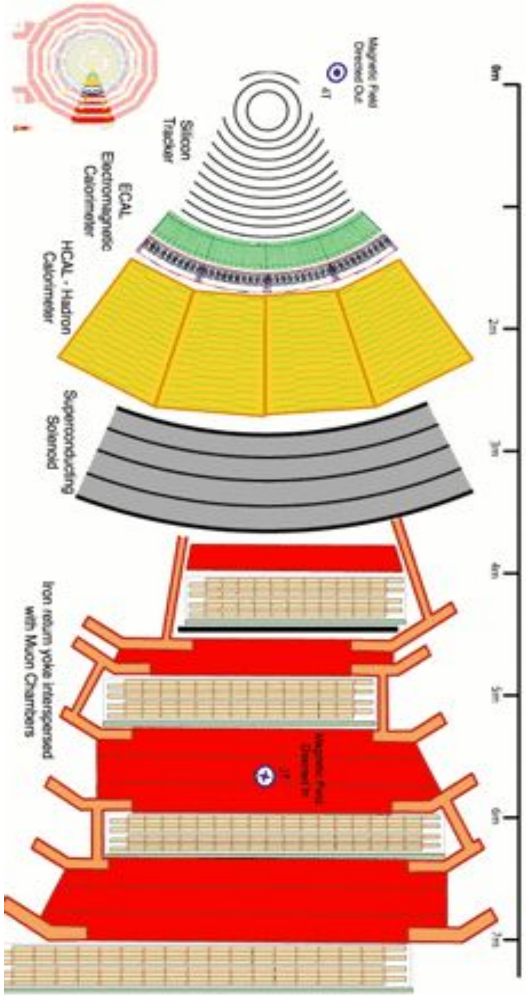
Jet Tagging 1.
Brief Introduction

ML for Jet Tagging 2.
State of the art

Our Proposal 3
The Boost Invariant Polynomials

Results 4.
Of the approach

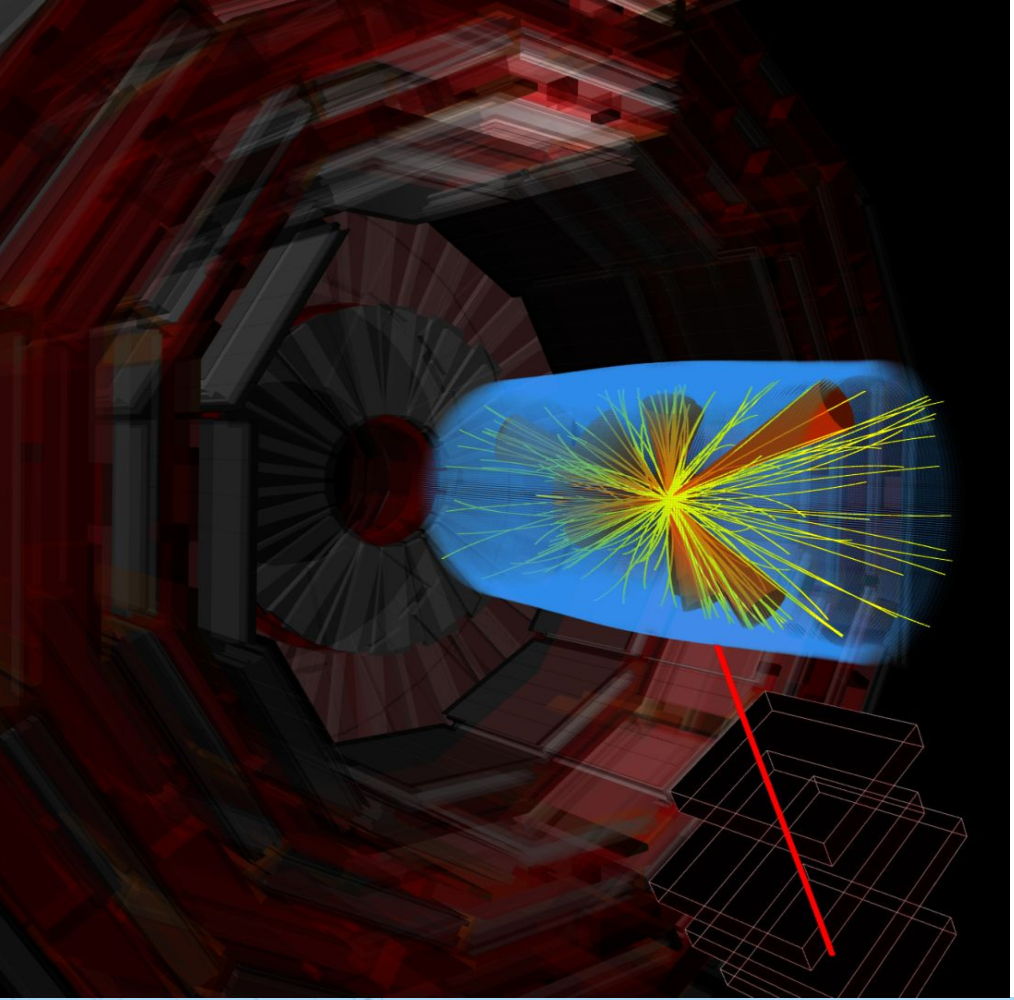
Future Outlook 5.
& Conclusions



1.

Jet Tagging

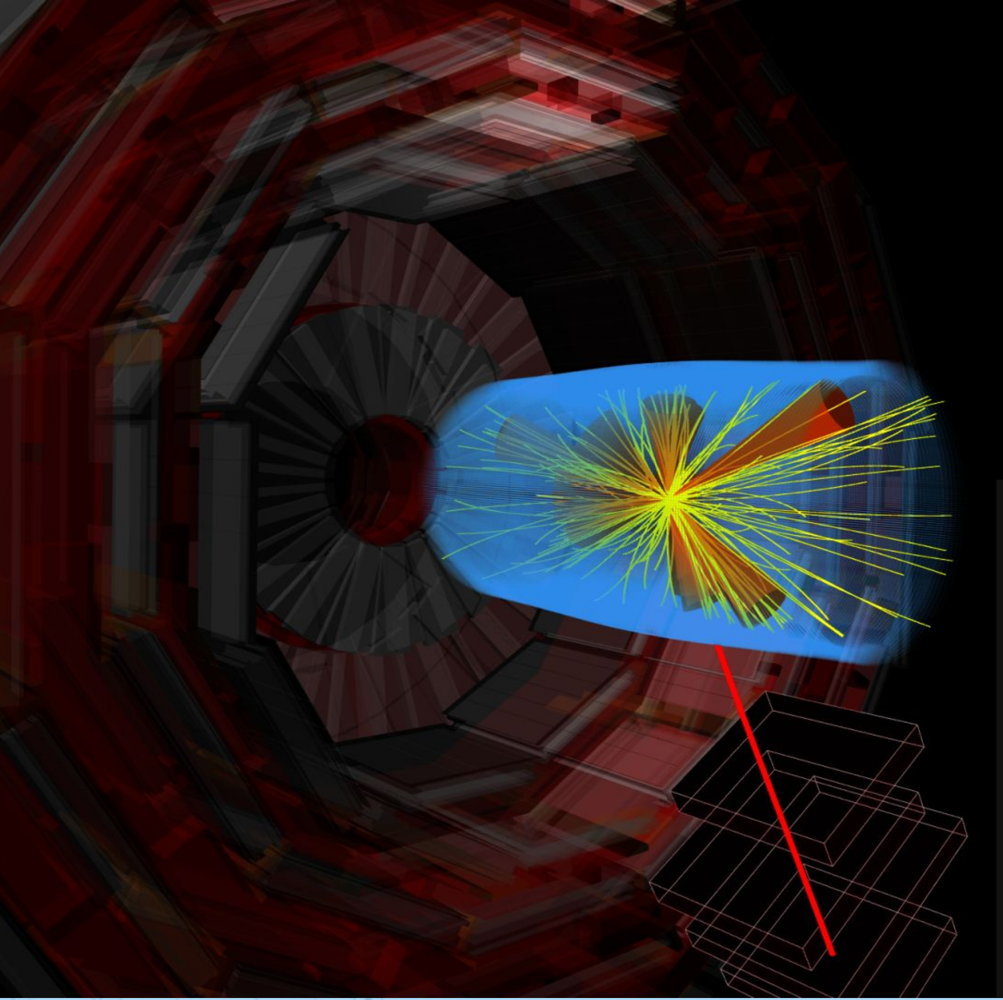
A Brief Introduction



Jet Tagging

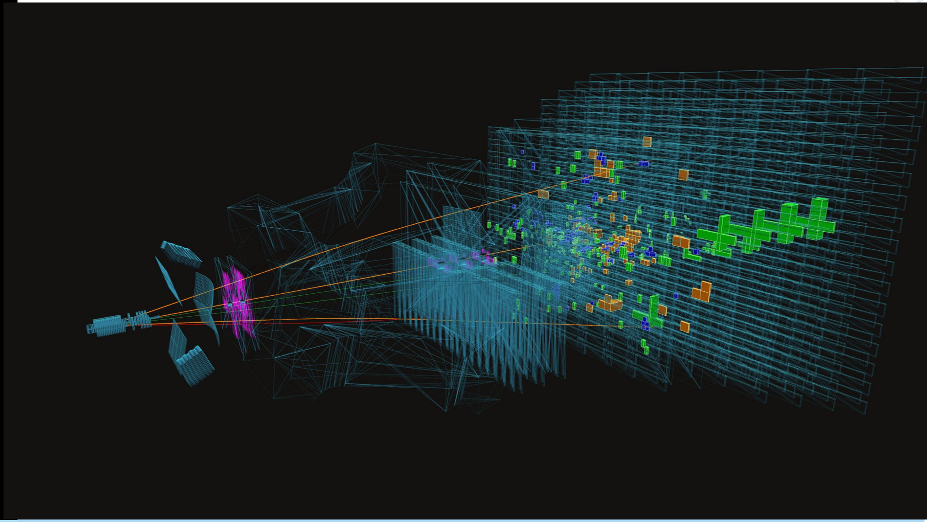
Identify signal vs background in particles coming from different processes

The particles are detected and reconstructed as point clouds



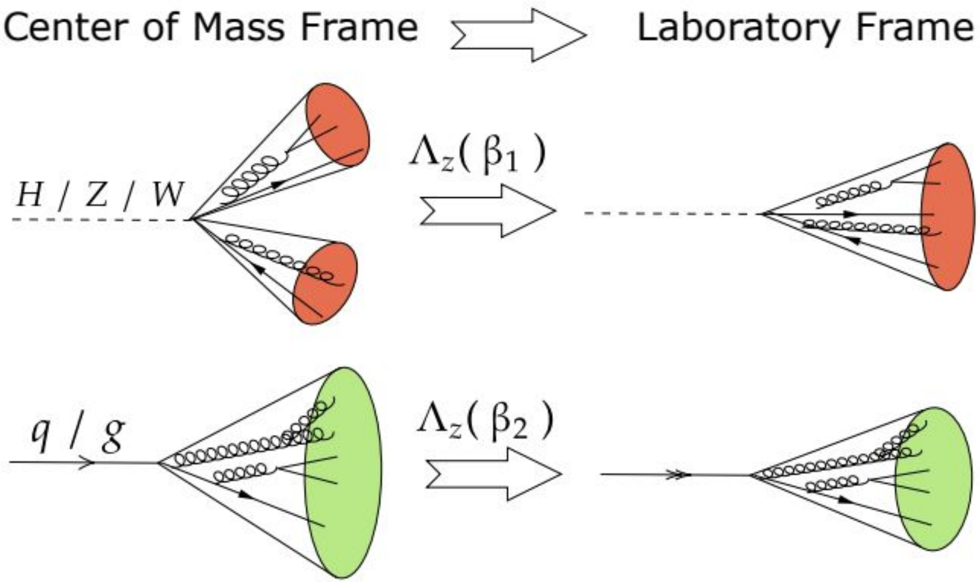
Jet Tagging

Identify signal vs background in particles coming from different processes



The issue of Boosted Jet Tagging

The angular resolution depends on the pT of the jets…



arXiv:2207.08272v2

2.

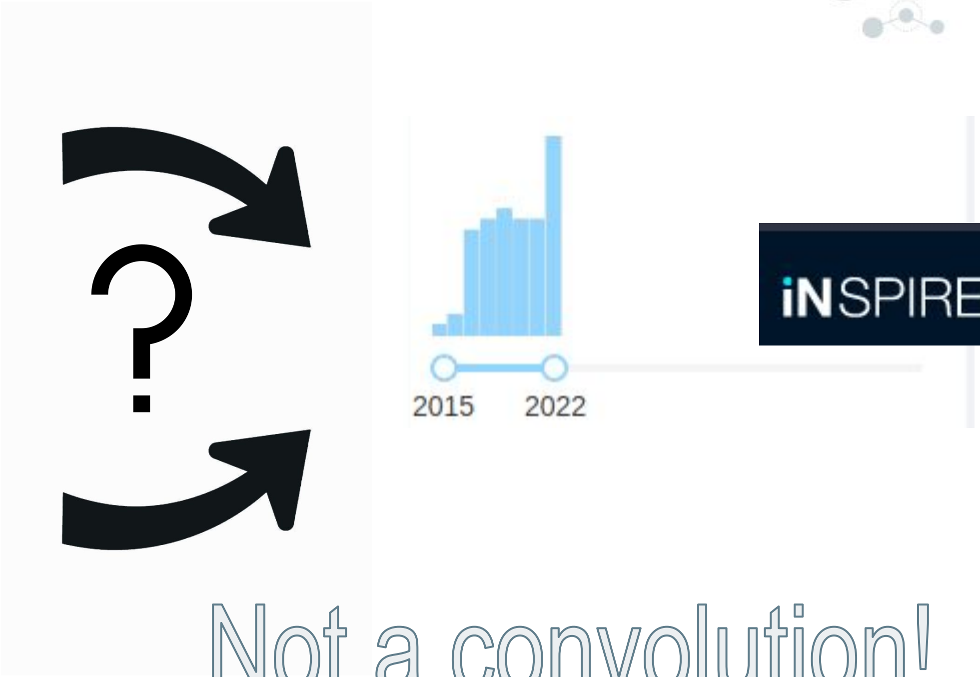
ML for Jet Tagging

A Brief Introduction

Why Machine Learning?

Jet-Tagging Papers:

ML papers:



Not a convolution!

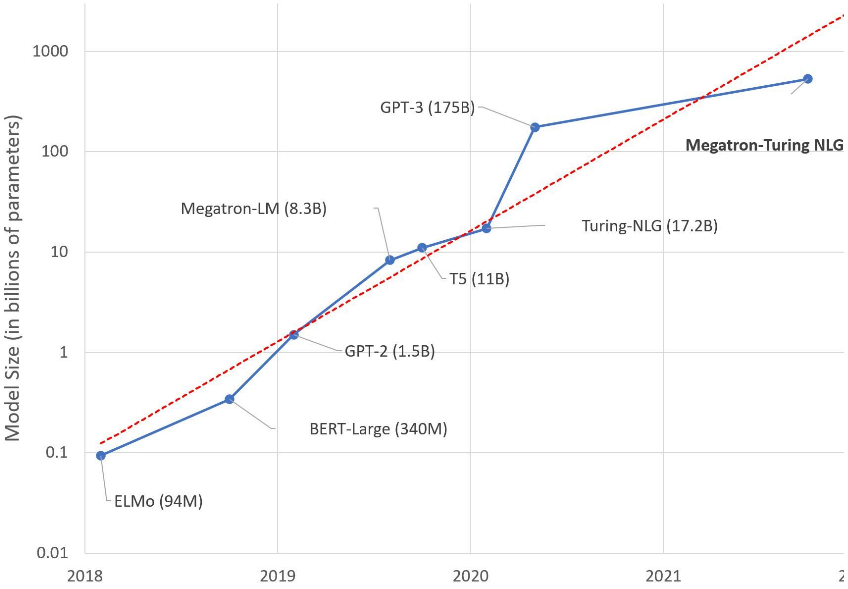
arXiv:2207.08272v2

The model size is a determining factor

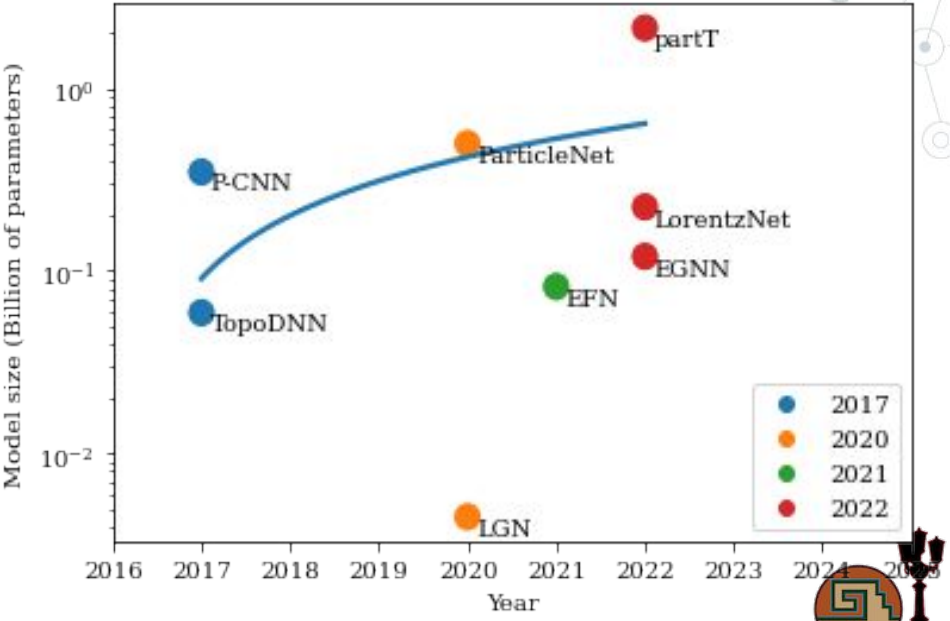
In other fields: (LLMs)

Particle physics

arXiv:2207.08272v2



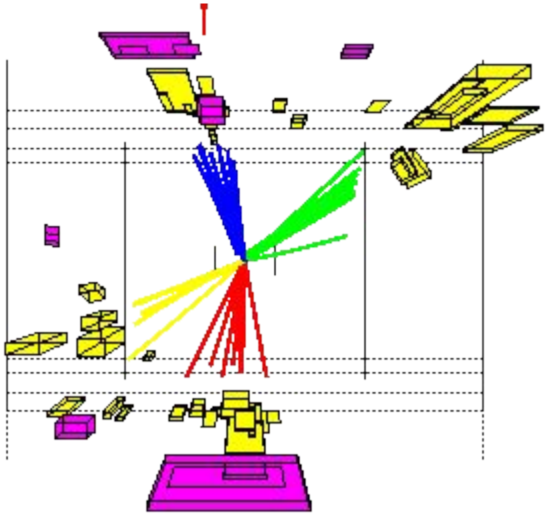
https://huggingface.co/blog/large-language-models



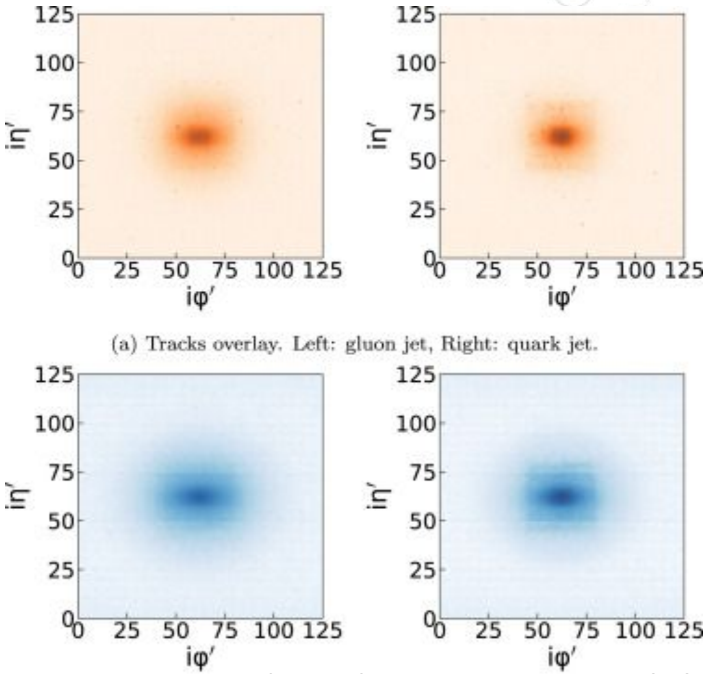
What information should we give the models?

arXiv:2207.08272v2

Low level variables

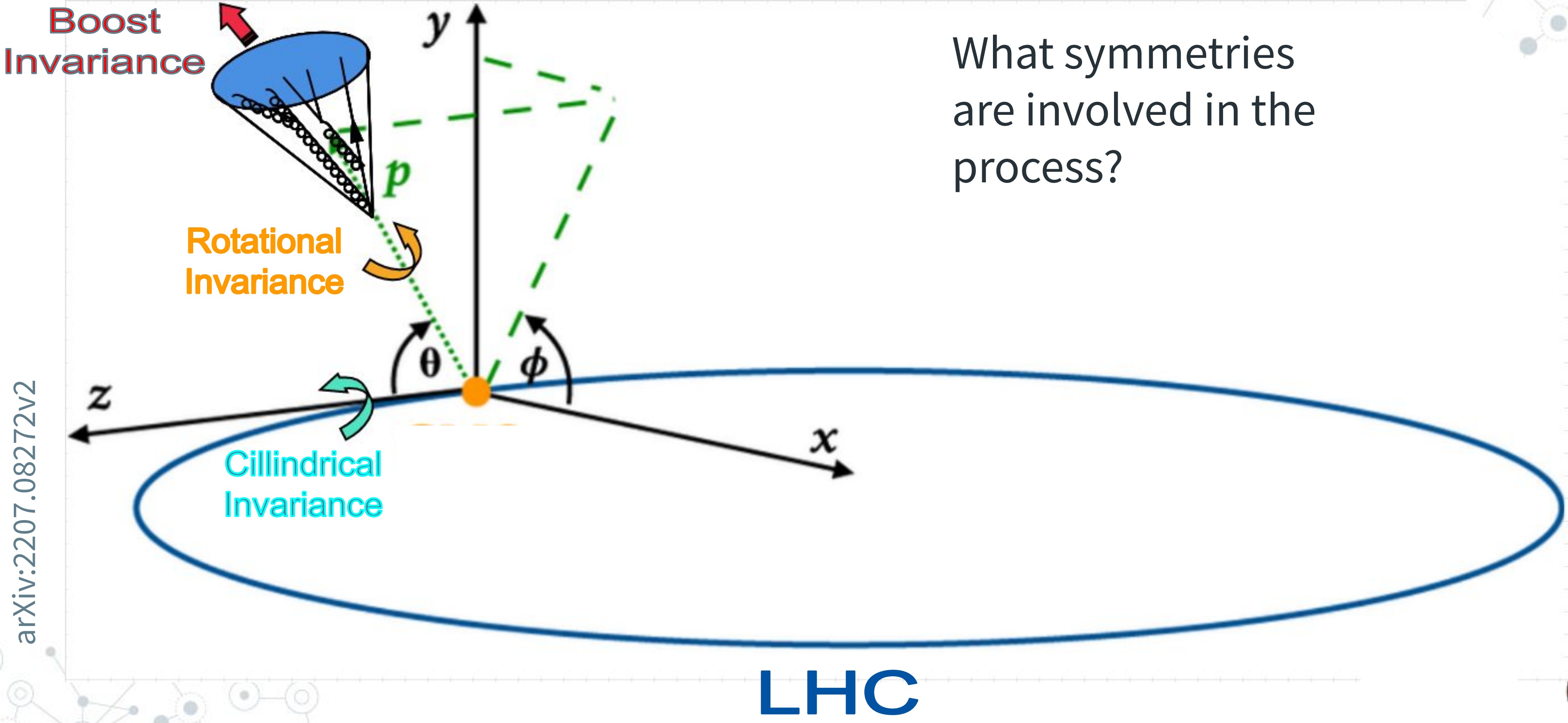


High Level variables



End-to-end jet classification of quarks and gluons with the CMS Open Data. Andrew et Al.

What about inductive biases?



arXiv:2207.08272v2

What symmetries are involved in the process?

What Lorentz group equivariant message passing networks achieve?

Equivariant Energy Flow Networks for jet tagging

Matthew J. Dolan[1,*] and Ayodele Ore[1,†]
[1] *ARC Centre of Excellence for Dark Matter Particle Physics, School of Physics, The University of Melbourne, Victoria 3010, Australia*

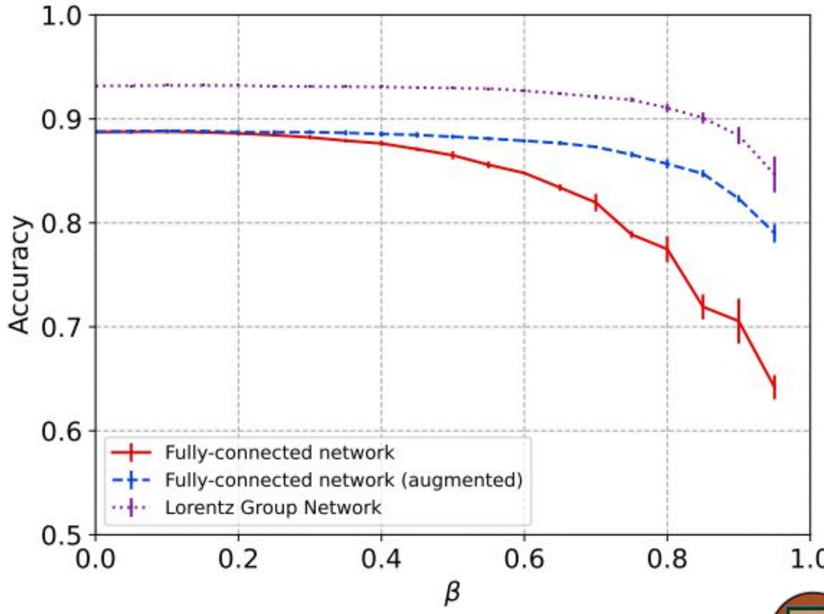
An Efficient Lorentz Equivariant Graph Neural Network for Jet Tagging

Shiqi Gong[a,e,1] Qi Meng[b] Jue Zhang[b] Huilin Qu[c] Congqiao Li[d] Sitian Qian[d] Weitao Du[a] Zhi-Ming Ma[a] Tie-Yan Liu[b]

arXiv:2207.08272v2

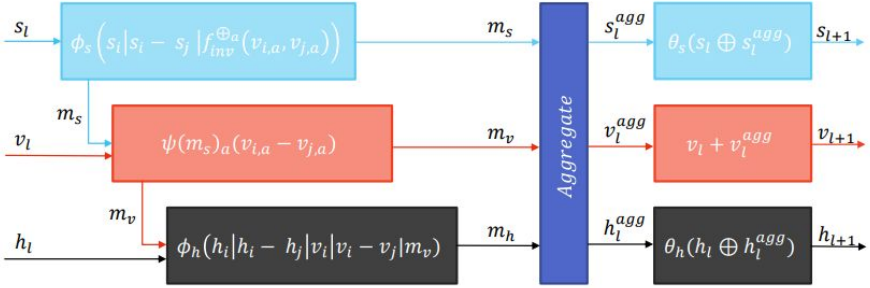
Semi-Equivariant GNN Architectures for Jet Tagging

Daniel Murnane[1], Savannah Thais[2], Jason Wong[3]

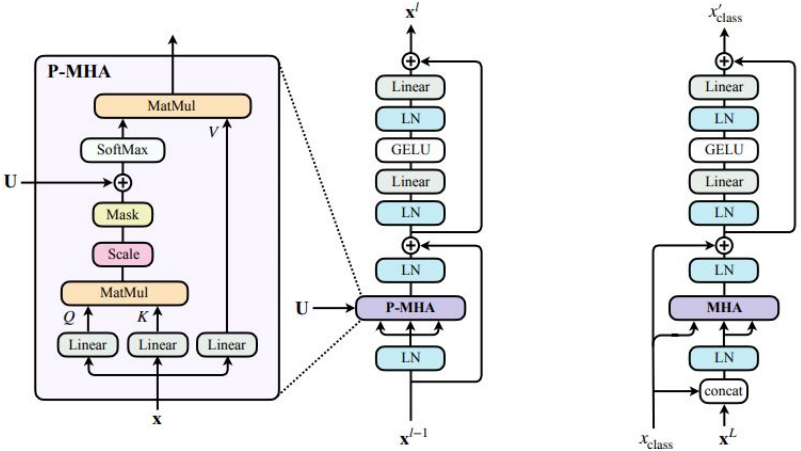


Bogatskiy et al.

A problem...

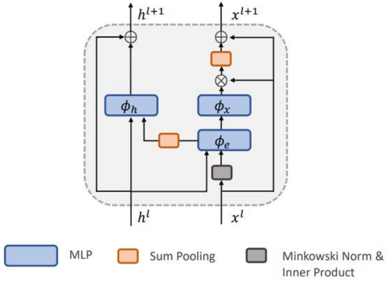


Murnane et al.

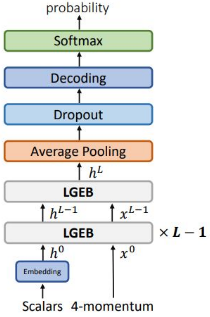


Qu et al.

arXiv:2207.08272v2



Lorentz Group Equivariant Block (LGEB)



LorentzNet

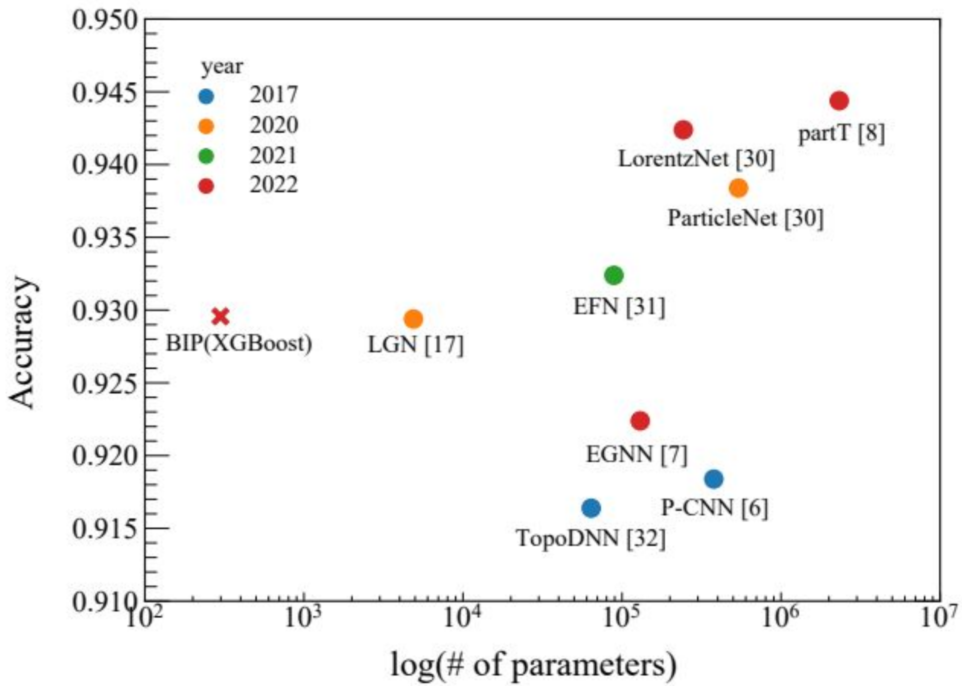
Very expensive tensorial computation

3.

The Boost Invariant Polynomials

A Brief Introduction

SOTA Accuracy with a fraction of the parameters



arXiv:2207.08272v2

A basis change

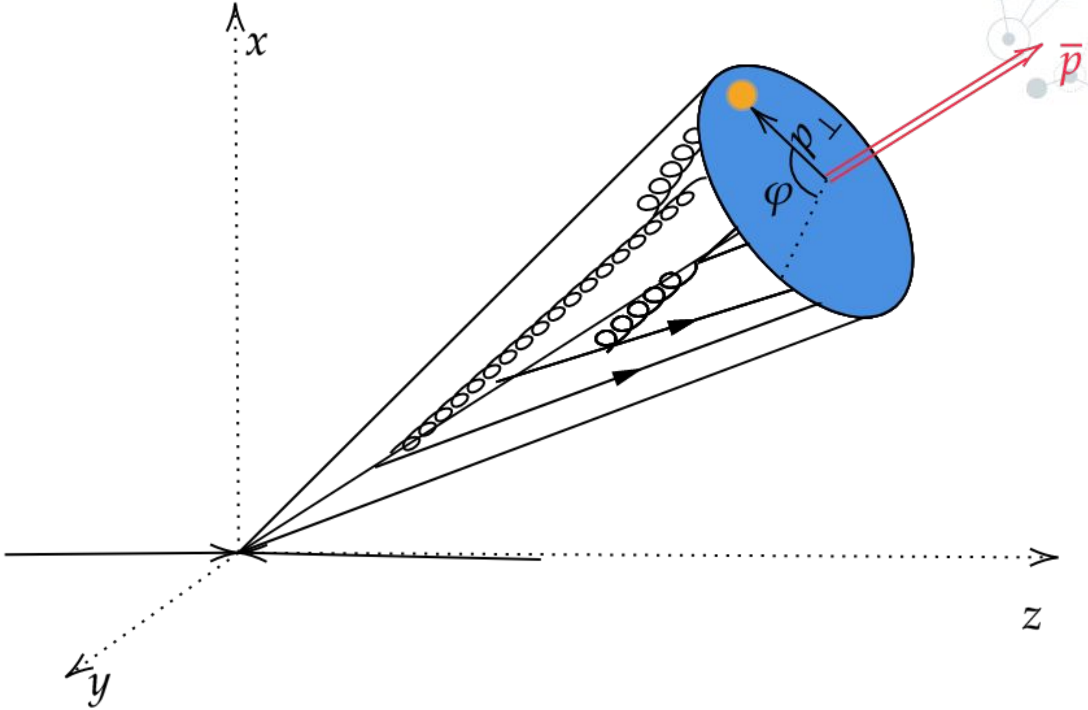
$$y_i = \frac{1}{2}\log\left(\frac{\delta_1 + E_i + p_{\|i}}{\delta_1 + E_i - p_{\|,i}}\right)$$

$$p_\perp$$

$$E_{\perp,i} = \sqrt{m_i^2 + p_{\perp,i}^2}$$

$$\varphi_i$$

arXiv:2207.08272v2



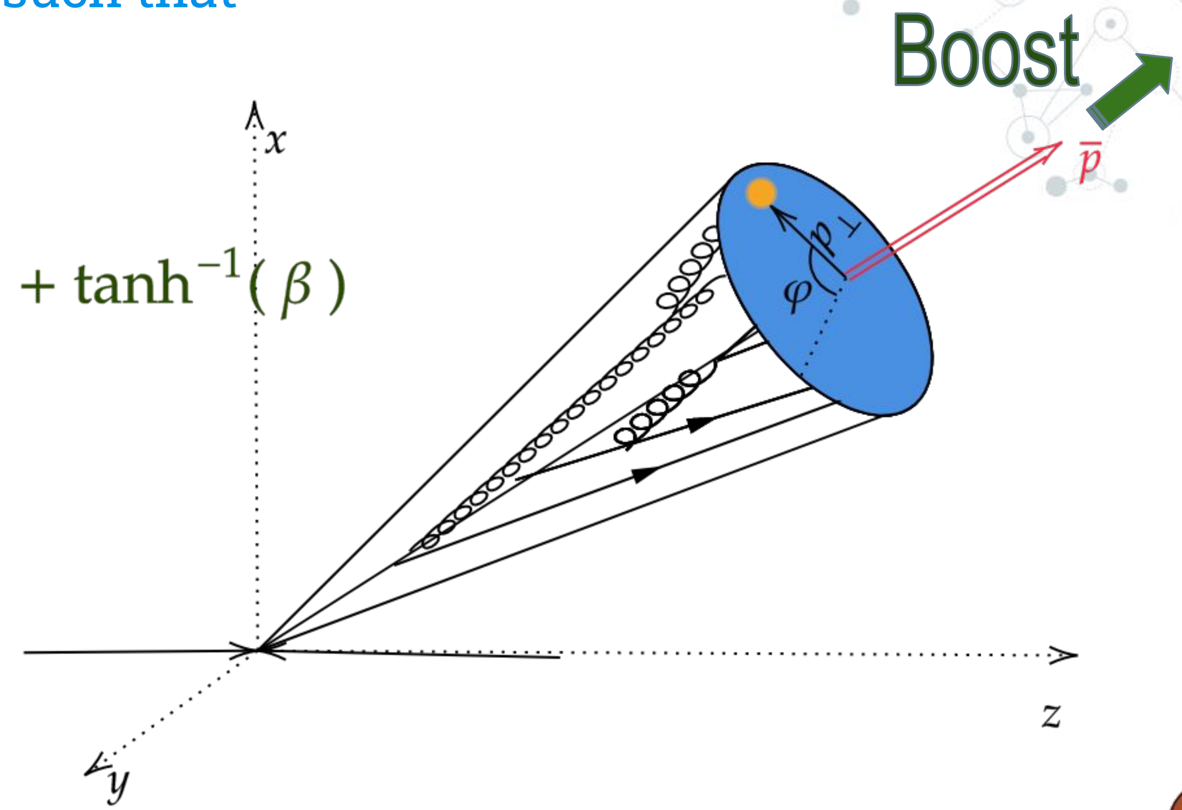
A basis change… such that

$$y_i = \frac{1}{2}\log\left(\frac{\delta_1 + E_i + p_{\|i}}{\delta_1 + E_i - p_{\|,i}}\right) + \tanh^{-1}(\,\beta\,)$$

$$p_\perp$$

$$E_{\perp,i} = \sqrt{m_i^2 + p_{\perp,i}^2}$$

$$\varphi_i$$



arXiv:2207.08272v2

A basis change… such that

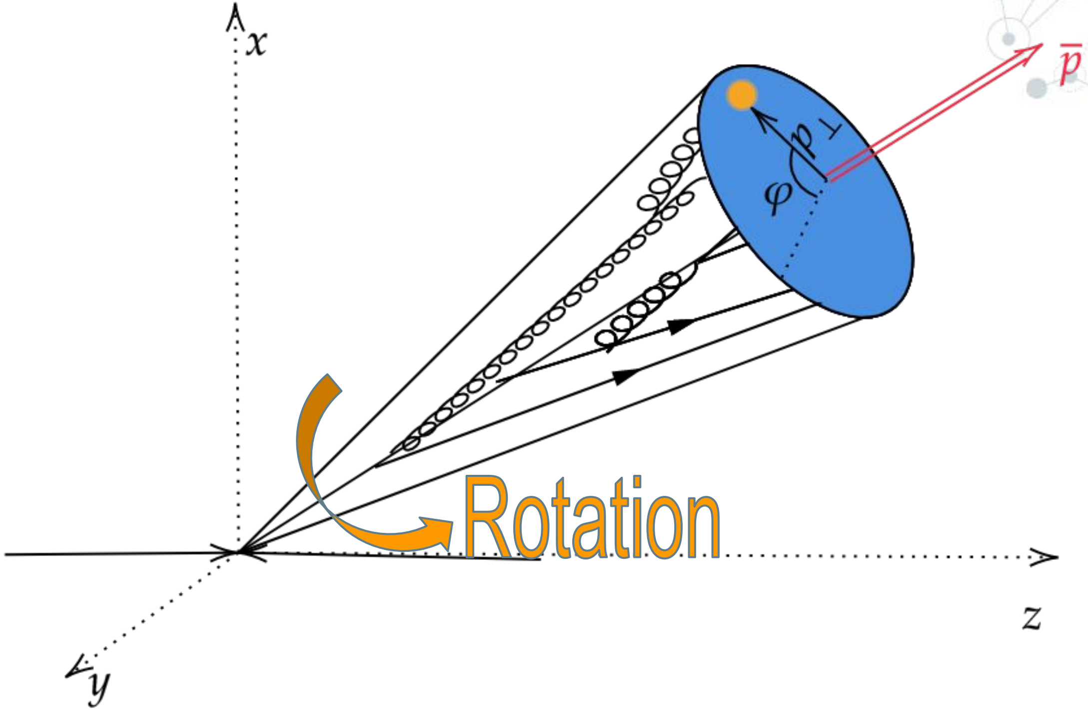
$$y_i = \frac{1}{2}\log\left(\frac{\delta_1 + E_i + p_{\|i}}{\delta_1 + E_i - p_{\|,i}}\right)$$

$$p_\perp$$

$$E_{\perp,i} = \sqrt{m_i^2 + p_{\perp,i}^2}$$

$$\varphi_i \quad + \Delta\varphi$$

arXiv:2207.08272v2



A basis change… such that

$$y_i = \frac{1}{2}\log\left(\frac{\delta_1 + E_i + p_{\|i}}{\delta_1 + E_i - p_{\|,i}}\right)$$

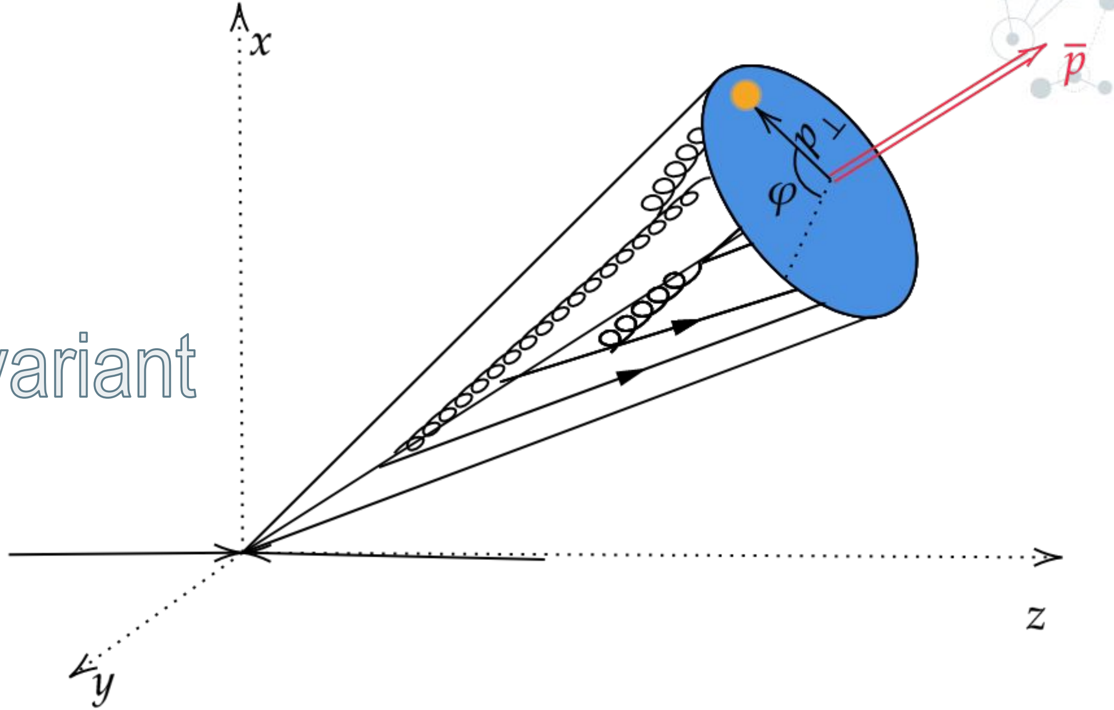
$$p_\perp$$

$$E_{\perp,i} = \sqrt{m_i^2 + p_{\perp,i}^2}$$

$$\varphi_i$$

Invariant

arXiv:2207.08272v2



The Atomic Cluster Expansion Recipe:

1. Build a one particle basis:

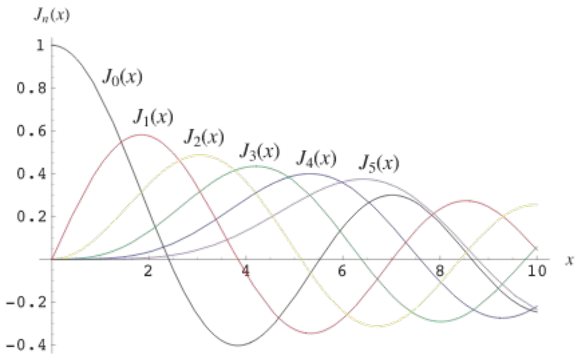
$$A_{nlk} = \sum_{i=1}^{N} Q_n(p_{\perp,i}, E_{\perp,i}, \xi_i) e^{il\varphi_i} e^{-\lambda k y_i}$$

The Atomic Cluster Expansion Recipe:

1. Build a one particle basis:

$$A_{nlk} = \sum_{i=1}^{N} Q_n(p_{\perp,i}, E_{\perp,i}, \xi_i) e^{il\varphi_i} e^{-\lambda k y_i}$$

arXiv:2207.08272v2



A general basis allows to introduce additional features… e.g

$$Q_n(E_{\perp,i}, p_{\perp,i}) = B_n(\tilde{p}_{\perp,i}) \log(1 + E_{\perp,i})$$

Bessel polynomials

The Atomic Cluster Expansion Recipe:

1. Build a one particle basis:

$$A_{nlk} = \sum_{i=1}^{N} Q_n(p_{\perp,i}, E_{\perp,i}, \xi_i) e^{il\varphi_i} e^{-\lambda k y_i}$$

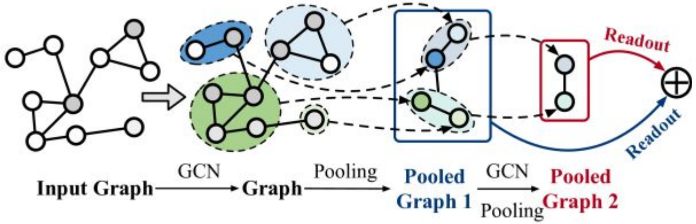
Decoupled generators
(rotations + boosts)

arXiv:2207.08272v2

The Atomic Cluster Expansion Recipe:

1. Build a one particle basis:

$$A_{nlk} = \sum_{i=1}^{N} Q_n(p_{\perp,i}, E_{\perp,i}, \xi_i) e^{\mathrm{i}l\varphi_i} e^{-\lambda k y_i}$$



Liu et al.

arXiv:2207.08272v2

The Atomic Cluster Expansion Recipe:

1. Build a one particle basis:

$$A_{nlk} = \sum_{i=1}^{N} Q_n(p_{\perp,i}, E_{\perp,i}, \xi_i) e^{il\varphi_i} e^{-\lambda k y_i}$$

NOT an Invariant jet descriptor

arXiv:2207.08272v2

The Atomic Cluster Expansion Recipe:

1. Build a one particle basis.

2. Symmetrize the basis:

$$\boldsymbol{A_{nlk}} = \prod_{t=1}^{\nu} A_{n_t l_t k_t}$$

$$(n_1 l_1 k_1,\ \dots,\ n_\nu l_\nu k_\nu)$$

The Atomic Cluster Expansion Recipe:

1. Build a one particle basis.

2. Symmetrize the basis.

3. Keep Invariant features:

$$\boldsymbol{A_{nlk}} = \prod_{t=1}^{\nu} \sum_{i=1}^{N} Q_n(p_{\perp,i}, E_{\perp,i}, \xi_i) e^{\mathrm{i} l \varphi_i} e^{-\lambda k y_i} \Rightarrow \begin{array}{l} \sum_t l_t = \sum_t k_t = 0 \\ \sum_{t=1}^{\nu} |l_t| + |k_t| + n_t \leq \Gamma \end{array}$$

The Atomic Cluster Expansion Recipe:

1. Build a one particle basis.
2. Symmetrize the basis.
3. Keep Invariant features.

et Voilà

A highly expressive and invariant jet representation

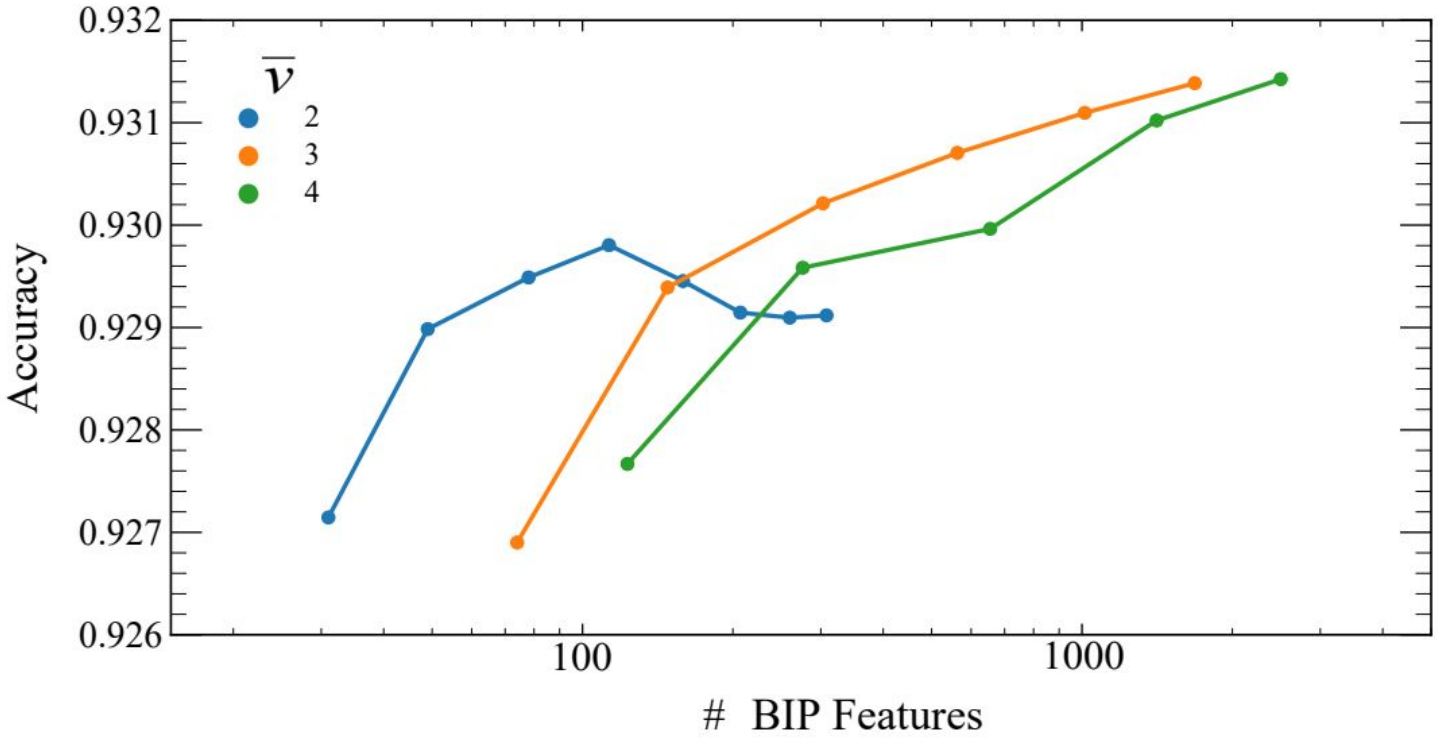
arXiv:2207.08272v2

4.

Results

Of the approach

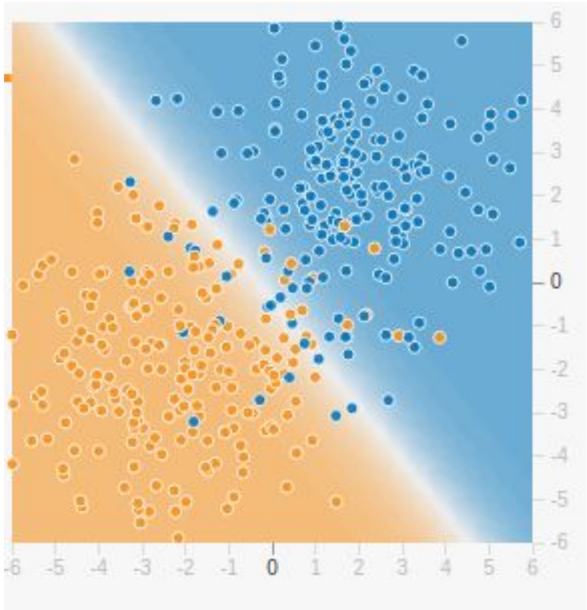
Versatility on the basis size



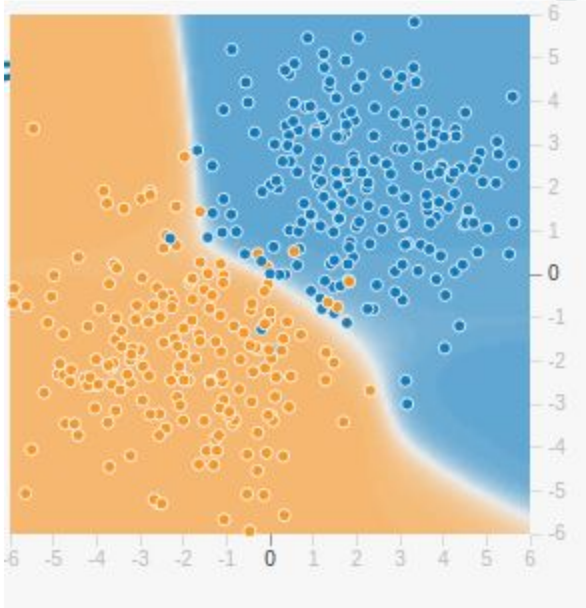
arXiv:2207.08272v2

Versatility on the classifier algorithm

From logistic regressions



To DL and GBT



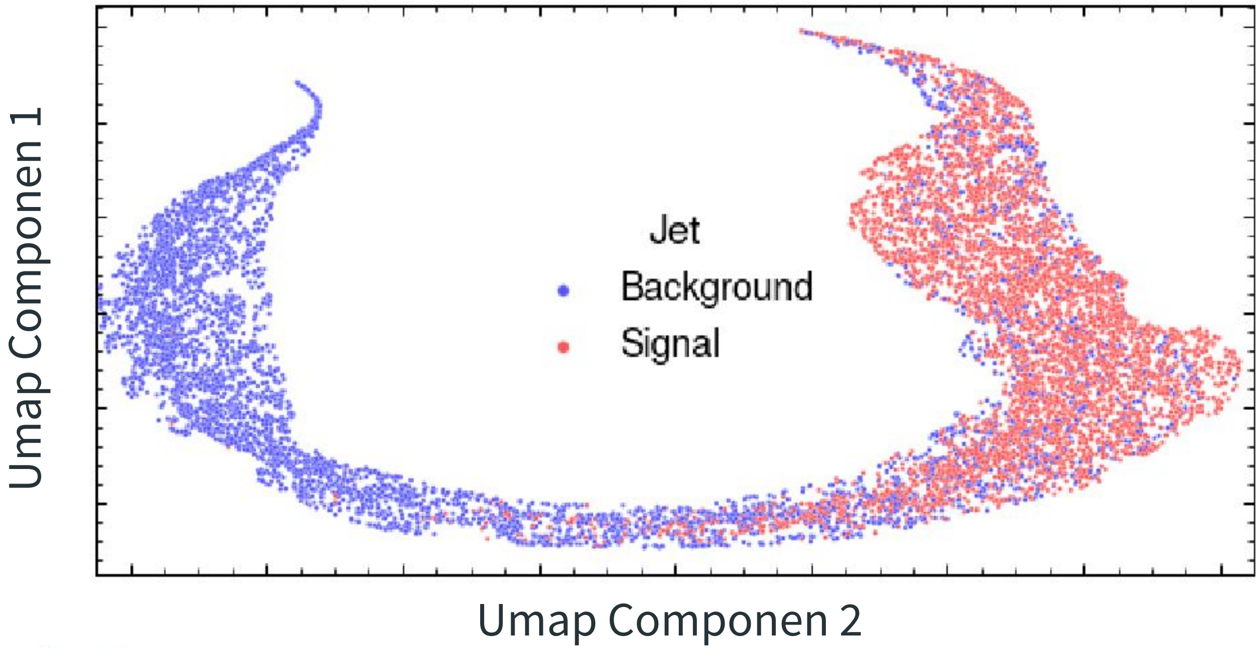
arXiv:2207.08272v2

Versatility on the classifier algorithm

Architecture	#Param	Accuracy	AUC
***partT (2022)**[8]	2.14M	0.944	0.988
EGNN (2022)[7]	120k	0.922	0.970
PCT (2021)[30]	139.3k	0.940	0.986
EFN (2021)[31]	82k	0.927	0.979
ParticleNet (2020)[30]	498k	0.938	0.985
LGN (2020)[17]	4.5k	0.929	0.964
P-CNN (2017)[6]	348k	0.918	0.980
TopoDNN (2017)[32]	59k	0.916	0.972
Supervised			
BIP(3, 6, MLP)	4k	0.931	0.981
BIP(3, 6, XGBoost)	300	0.929	0.978
BIP(3, 6, LogReg)	300	0.927	0.977
BIP(3, 6, SVM)	300	0.927	0.976

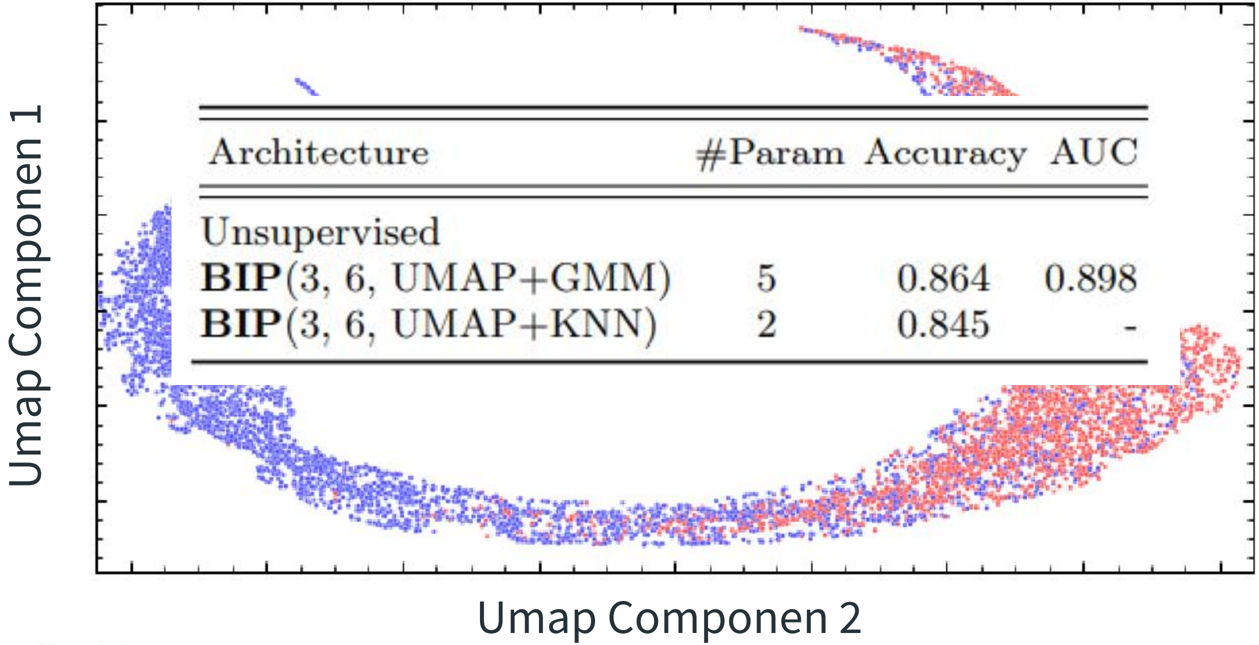
arXiv:2207.08272v2

Expressive representation



arXiv:2207.08272v2

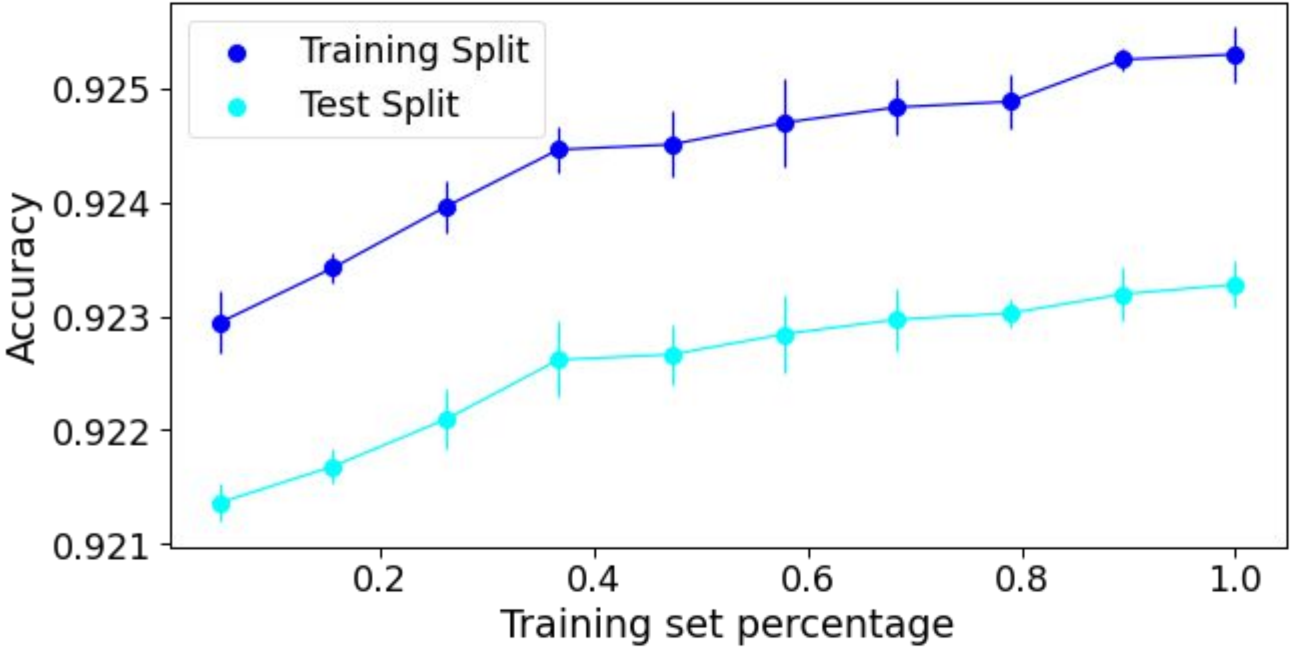
Expressive representation



Architecture	#Param	Accuracy	AUC
Unsupervised			
BIP(3, 6, UMAP+GMM)	5	0.864	0.898
BIP(3, 6, UMAP+KNN)	2	0.845	-

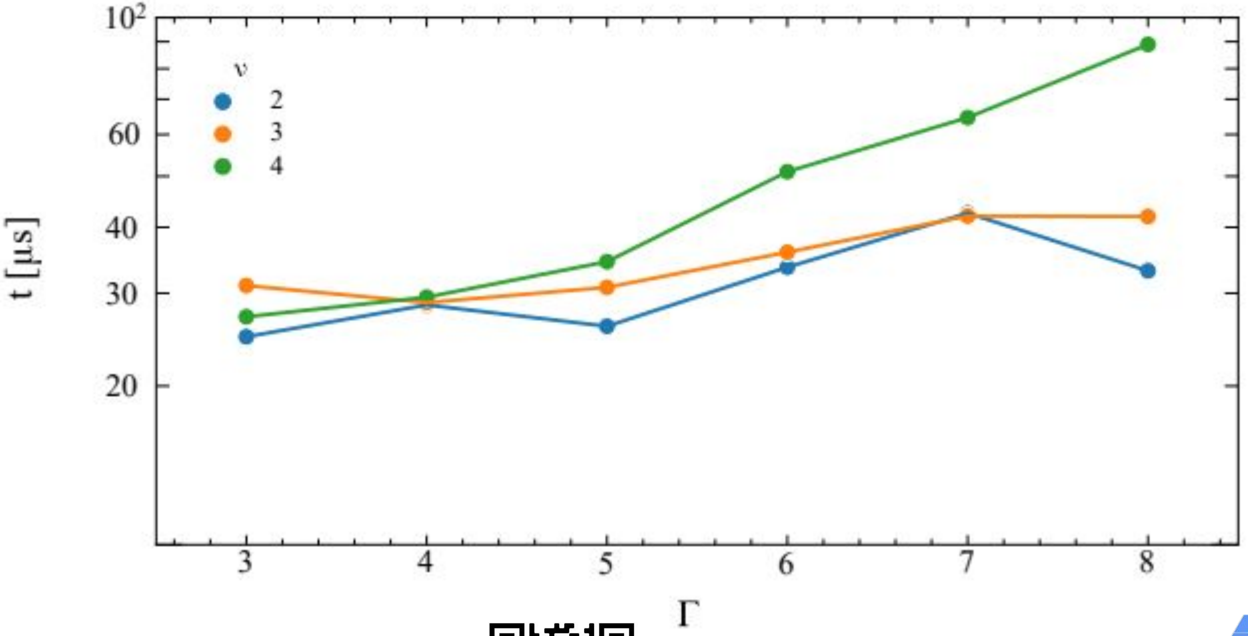
arXiv:2207.08272v2

Data Efficiency



arXiv:2207.08272v2

High efficiency



Also on…

same performance

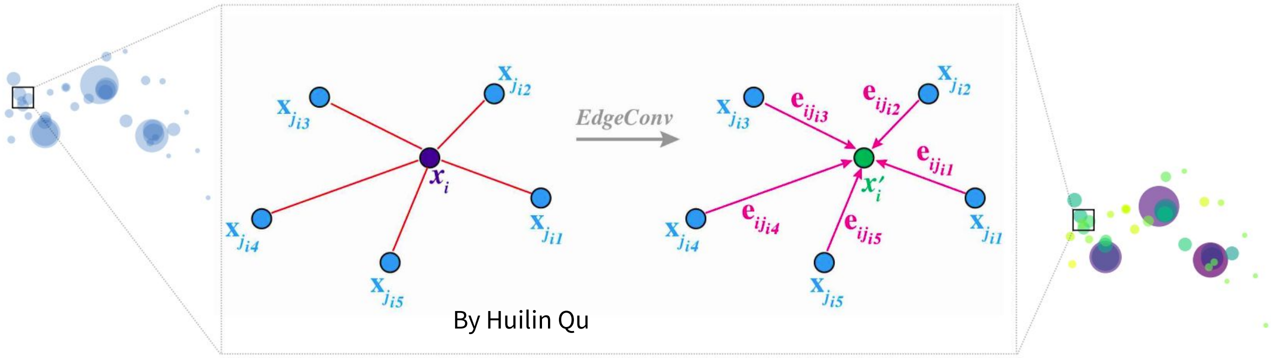
arXiv:2207.08272v2

5.

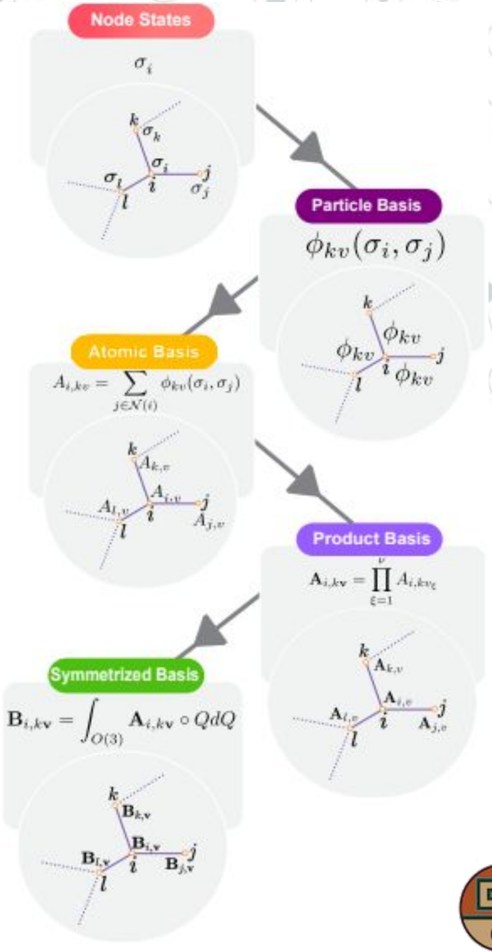
Outlooks

& Conclussions

Now... go deep



By Huilin Qu



arXiv:2207.08272v2

+ Learnable embeddings

Batatia et. al

Play with the model

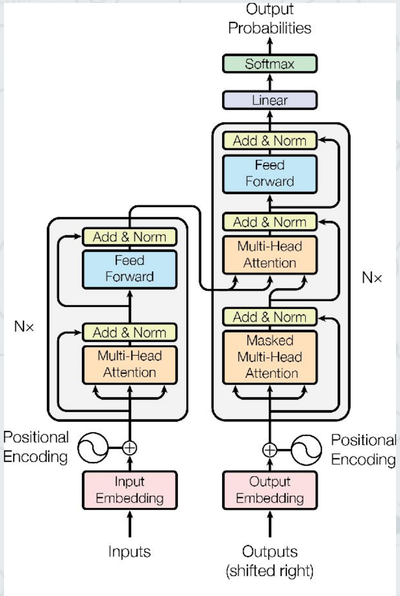
Paper

arXiv:2207.08272v2

Thanks for the ~~attention~~!

profile

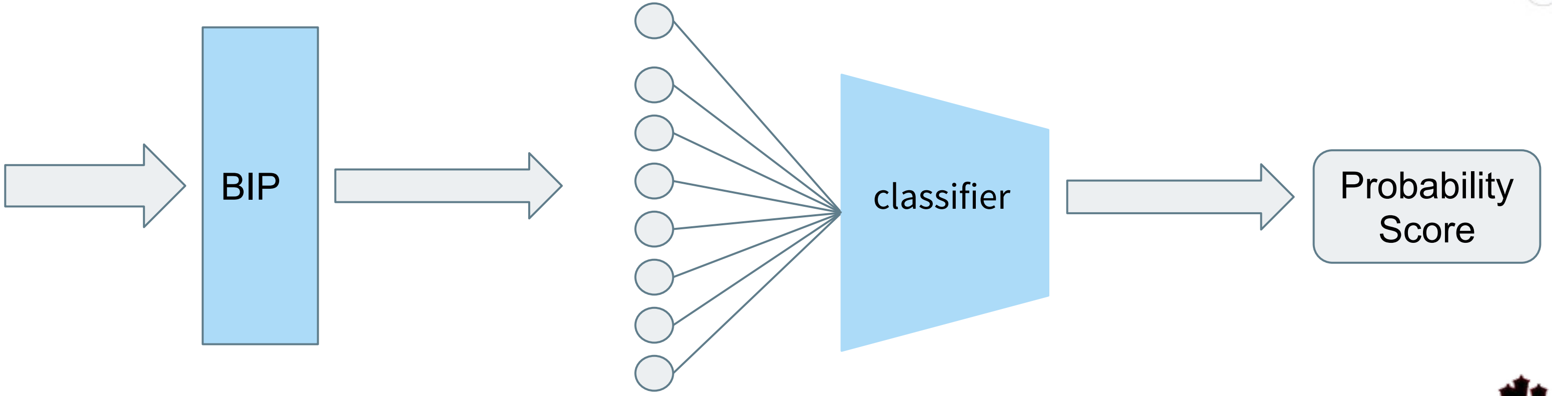
Please contact me for anything related




BACKUP

The classifier setup:

$$f(\{E_i, \boldsymbol{p}_i, \xi_i\}_i) = \sum_{\boldsymbol{nlk}} w_{\boldsymbol{nlk}} \boldsymbol{A}_{\boldsymbol{nlk}},$$



Jet representation

arXiv:2207.08272v2

A many body expansion

$$f(\{x_i\}_i) = f_0 + \sum_i f_1(x_i) + \sum_{i_1,i_2} f_2(x_{i_1}, x_{i_2}) + \cdots + \sum_{i_1,\ldots,i_{\bar{\nu}}} f_{\bar{\nu}}(x_{i_1}, \ldots, x_{i_{\bar{\nu}}}).$$

Or a density

$$\rho(x) = \sum_{i=1}^{N} \delta(x - x_i)$$

$$A_v = \langle Q_n(p_\perp, E_\perp, \xi) e^{il\varphi} e^{-ky} \mid \rho \rangle$$

$$\langle \phi_{v_1} \otimes \cdots \otimes \phi_{v_\nu} \mid \rho \otimes \cdots \otimes \rho \rangle = \prod_t \langle \phi_{v_t} \mid \rho \rangle = \prod_t A_{v_t} = \boldsymbol{A}_{\boldsymbol{v}},$$

arXiv:2207.08272v2