AI-assisted Four Top Quark Reconstruction
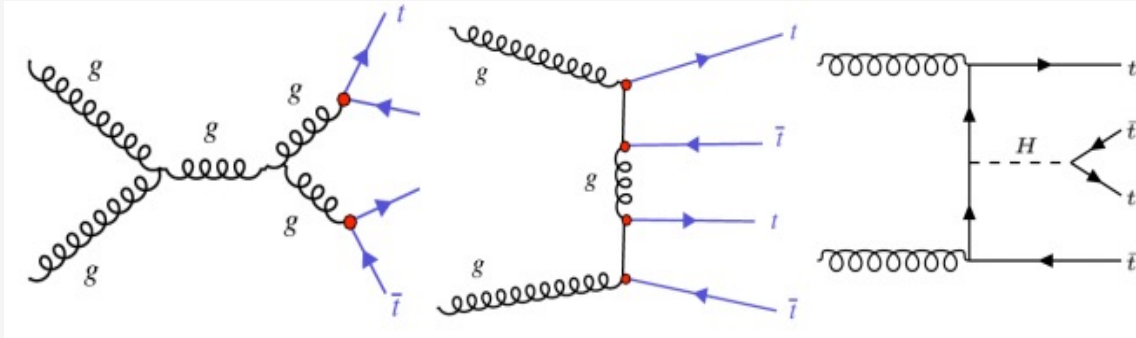
Xiang Chen, Yifan Wu, Jiahui Wu, Liang Li

Shanghai Jiao Tong University

2025/08/22

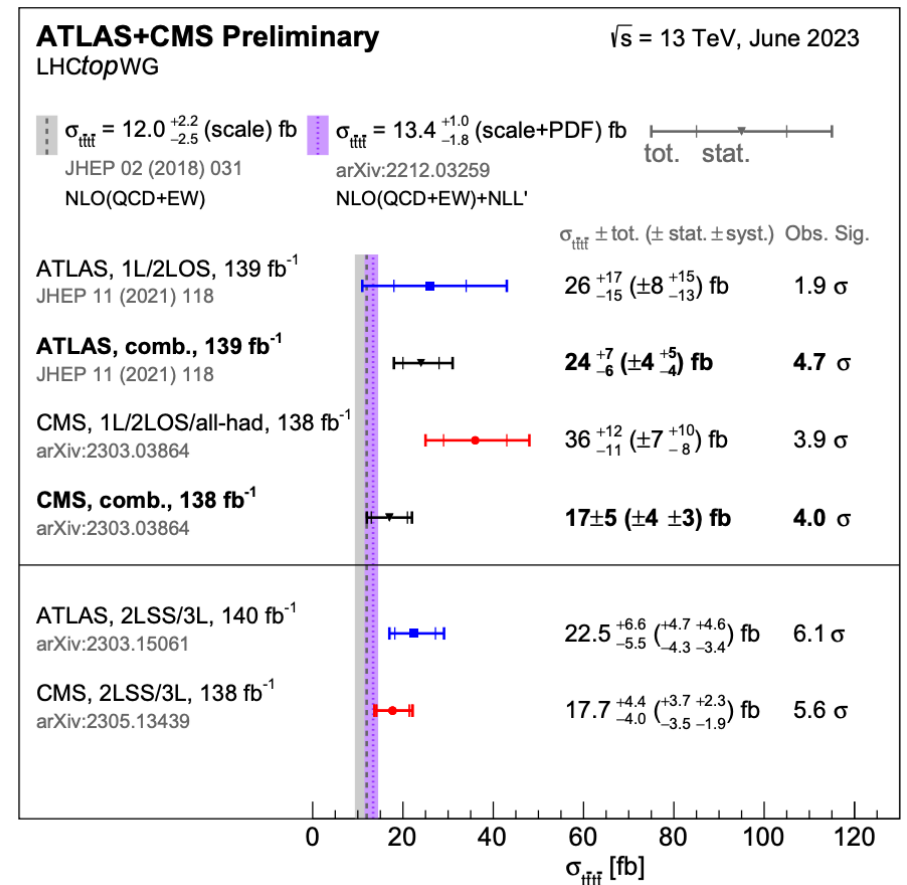# Rare Top Quark Production Process: Four-top-quark production

- $t\bar{t}t\bar{t}$ production is a rare top quark process predicted in the SM It is one of the heaviest final states accessible at LHC
  - NLO (QCD+EWK): $\sigma(t\bar{t}t\bar{t})$ = 12 fb ±20% [JHEP 02 (2018) 031]
  - NLO+NLL: $\sigma(t\bar{t}t\bar{t})$ = 13.4 fb ±11% [arXiv:2212.03259]

The observed significance is **6.1** sigma

ATLAS+CMS Preliminary
LHC*top*WG
√s = 13 TeV, June 2023

$\sigma_{t\bar{t}t\bar{t}}$ = 12.0 $^{+2.2}_{-2.5}$ (scale) fb
JHEP 02 (2018) 031
NLO(QCD+EW)

$\sigma_{t\bar{t}t\bar{t}}$ = 13.4 $^{+1.0}_{-1.8}$ (scale+PDF) fb
arXiv:2212.03259
NLO(QCD+EW)+NLL'

tot.   stat.

	$\sigma_{t\bar{t}t\bar{t}}$ ± tot. (± stat. ±syst.)	Obs. Sig.
ATLAS, 1L/2LOS, 139 fb⁻¹ JHEP 11 (2021) 118	26 $^{+17}_{-15}$ (±8 $^{+15}_{-13}$) fb	1.9 σ
ATLAS, comb., 139 fb⁻¹ JHEP 11 (2021) 118	24 $^{+7}_{-6}$ (±4 $^{+5}_{-4}$) fb	4.7 σ
CMS, 1L/2LOS/all-had, 138 fb⁻¹ arXiv:2303.03864	36 $^{+12}_{-11}$ (±7 $^{+10}_{-8}$) fb	3.9 σ
CMS, comb., 138 fb⁻¹ arXiv:2303.03864	17±5 (±4 ±3) fb	4.0 σ
ATLAS, 2LSS/3L, 140 fb⁻¹ arXiv:2303.15061	22.5 $^{+6.6}_{-5.5}$ ($^{+4.7}_{-4.3}$ $^{+4.6}_{-3.4}$) fb	6.1 σ
CMS, 2LSS/3L, 138 fb⁻¹ arXiv:2305.13439	17.7 $^{+4.4}_{-4.0}$ ($^{+3.7}_{-3.5}$ $^{+2.3}_{-1.9}$) fb	5.6 σ

0   20   40   60   80   100   120
$\sigma_{t\bar{t}t\bar{t}}$ [fb]

- $t\bar{t}t\bar{t}$ cross section is sensitive to anomalous top Yukawa coupling and Higgs CP properties
- Most sensitive channels in ATLAS and CMS are lepton channels: **SSML** and **1LOS**

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

Top Quark Reconstruction

- Top quark reconstruction relies on the measurement of final states and jet-parton assignment is important in heavy particle reconstruction.

- Traditionally we use the χ^2 **method** to build each possible permutation of the event to find the best solution

  - Large number of permutations and long running time
  - Hard to reconstruct leptonic top
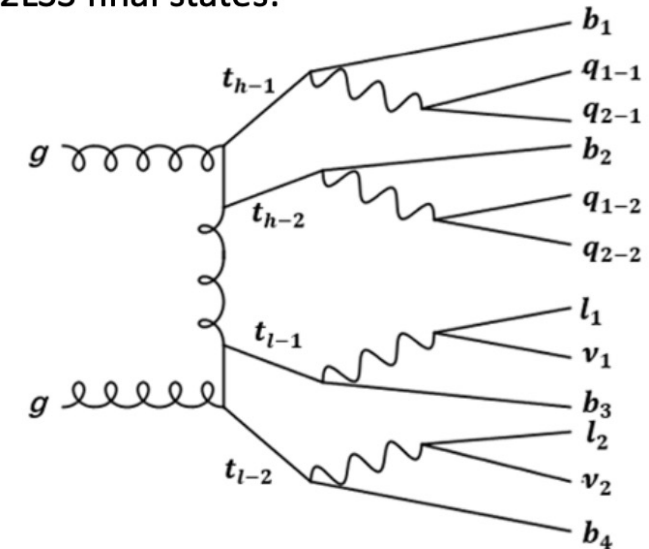- Depending on b-tagging

12 jets and 4 b jets: 2520 permutations
14 jets and 5 b jets: 113400 permutations

Find the best permutation with smallest χ^2

$$\chi^2 = \frac{(m_{b_1j_1j_2}-m_{top})^2}{\sigma_{top}^2} + \frac{(m_{b_2j_3j_4}-m_{top})^2}{\sigma_{top}^2} + \frac{(m_{j_1j_2}-m_W)^2}{\sigma_{m_W}^2} + \frac{(m_{j_3j_4}-m_W)^2}{\sigma_{m_W}^2}$$

Event Type	N jet	Event fraction	Hadronic top chi2 efficiency
All Event	==6	7.9%	0.07
	==7	19.8%	0.130
	>=8	72.9%	0.1985
	inclusive	1	0.172

2LSS final states:

Machine Learning Assignment Methods

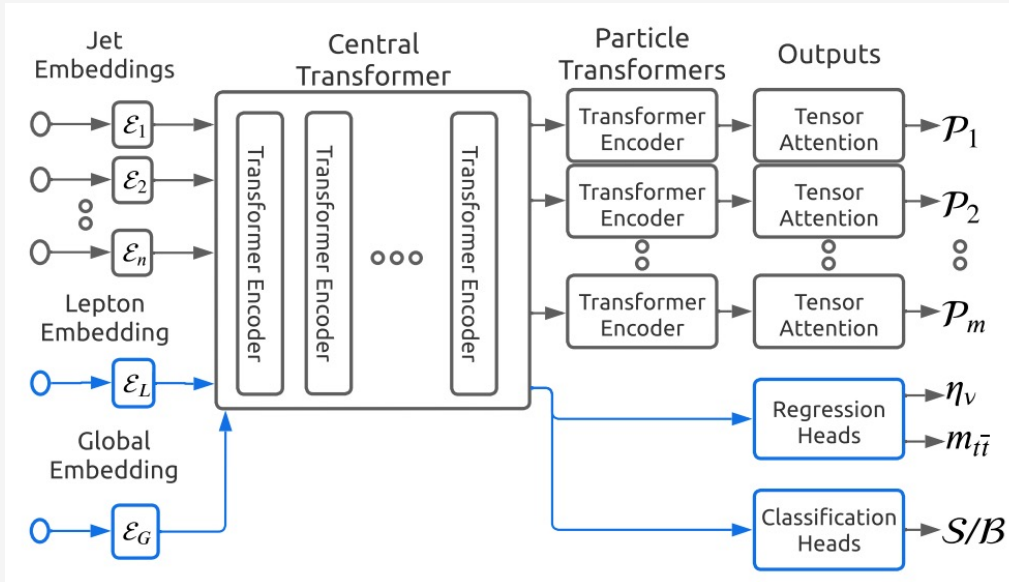Neural Network methods' advantages:

- Less time consumption in application: parallel computation by natural network
- High accuracy and can handle partial events

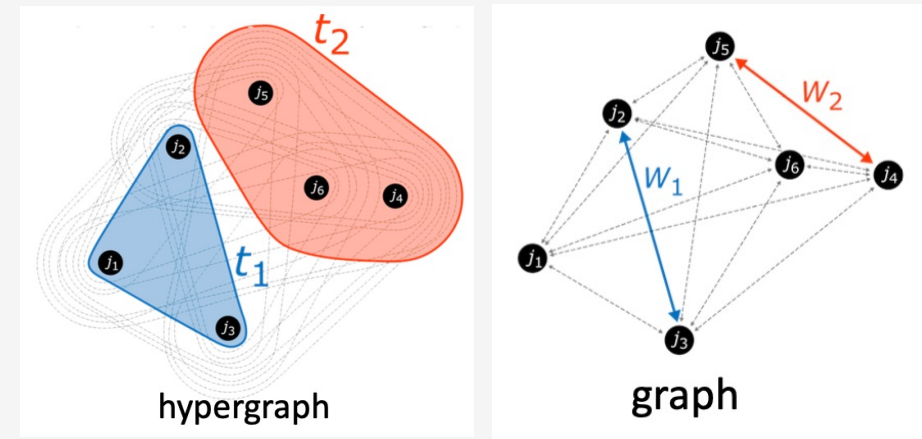Commun Phys **7**, 139 (2024)   SciPost Phys. 12, 178 (2022)

SPA-NET: use transformer to read multiple inputs from physics objects and output as sequence

$$Loss = \alpha_{reco}L_{reconstruction} + \alpha_{det}L_{detection} + \alpha_{reg}L_{regression}$$


PhysRevD.111.032004

Hyper: use hyper-graph to group the top quark candidates in the graph dataset

Truth Labelling for Supervised Learning

- Event generated by Madgraph5 and go trough fast simulation using Delphes

- Surprised learning: need to label the jet in the final state
    - **Spatial matching** based on truth-level partons in (eta, phi) plane
    - Match the nearest jet to each parton: labelling efficiency is ~75%

- ~600k 4top sample in 2L channel for training and validation

Dataset	Number of events
Training	551k
Validation	29k
Evaluation	175k

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\varphi^2} < 0.3$$
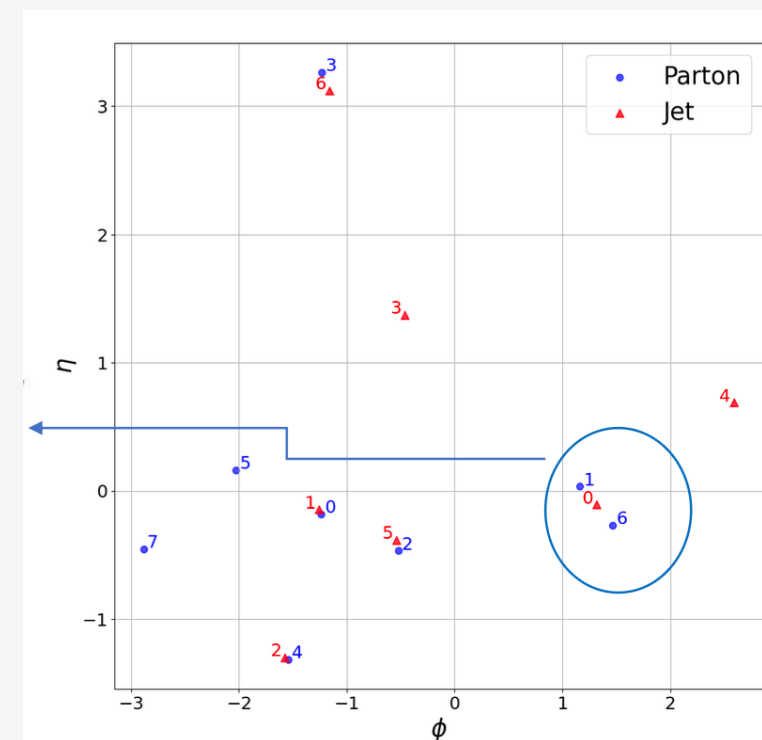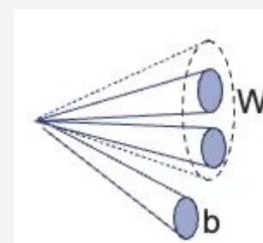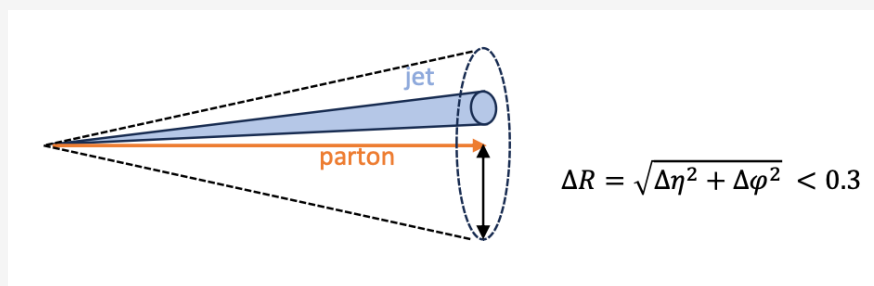
SPA-Net Performance in Four-Top-Quark Reconstruction

- Evaluating the performance using the top quarks which decay products are spatially matched to partons

- The machine learning method outperforms the traditional method in accuracy and speed:
  - Speed: SPA-Net improves about **3 times** faster in running time(both in CPU)
  - Accuracy: SPA-Net achieves **~38%** in hadronic top efficiency, **~70%** for leptonic top efficiency for four-top-quark events
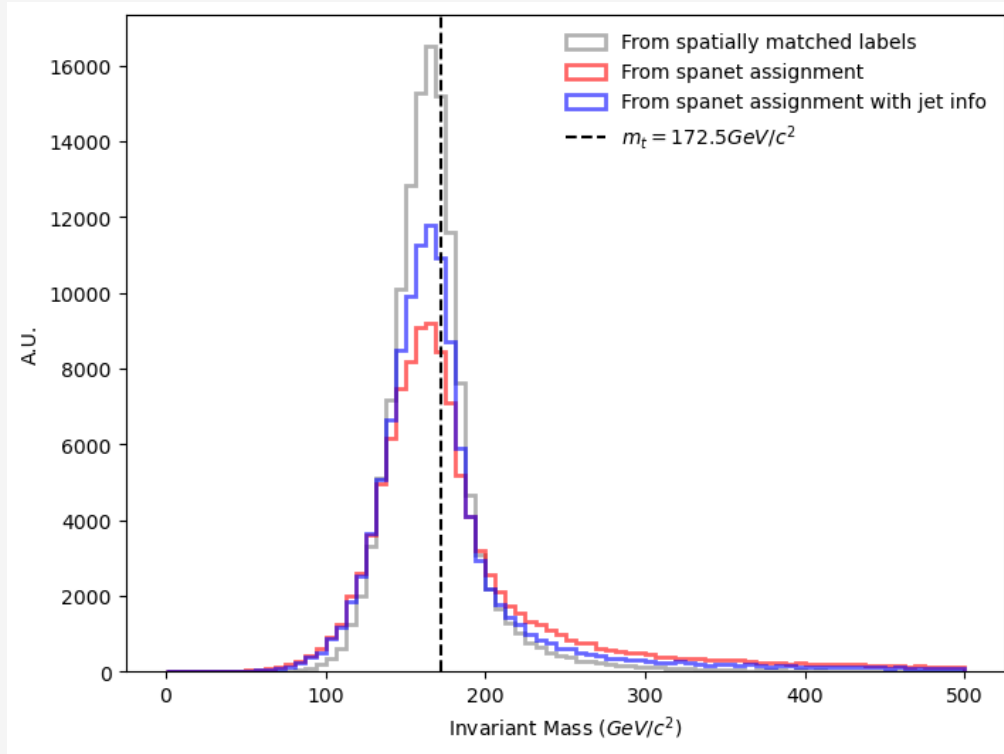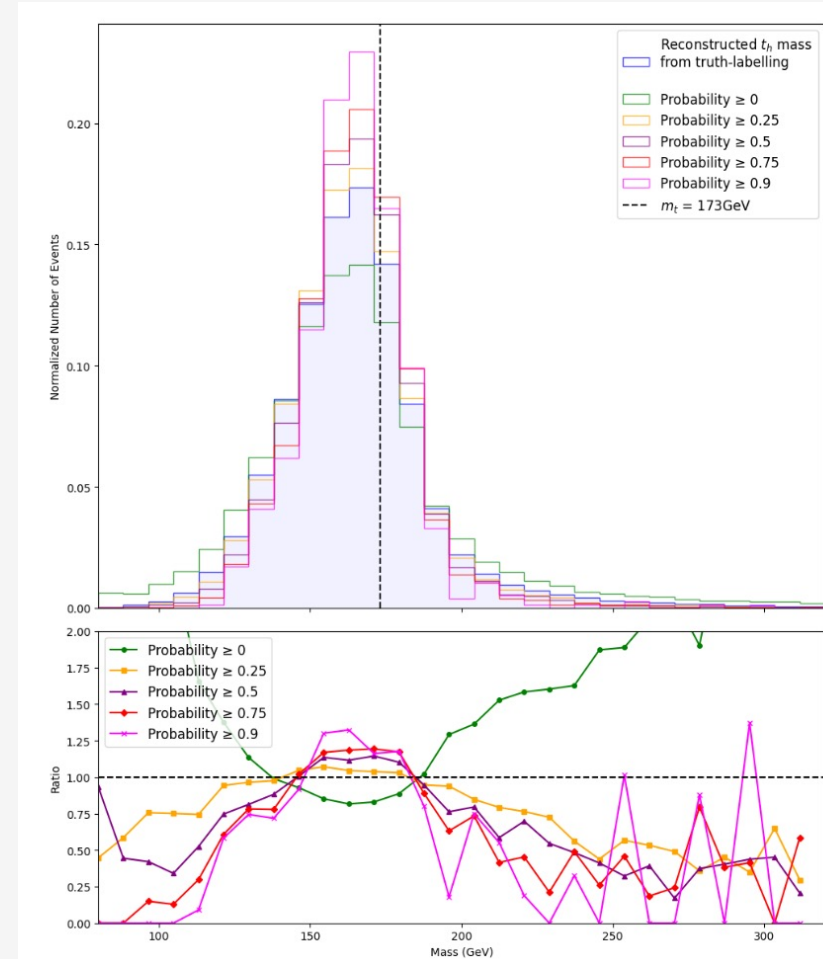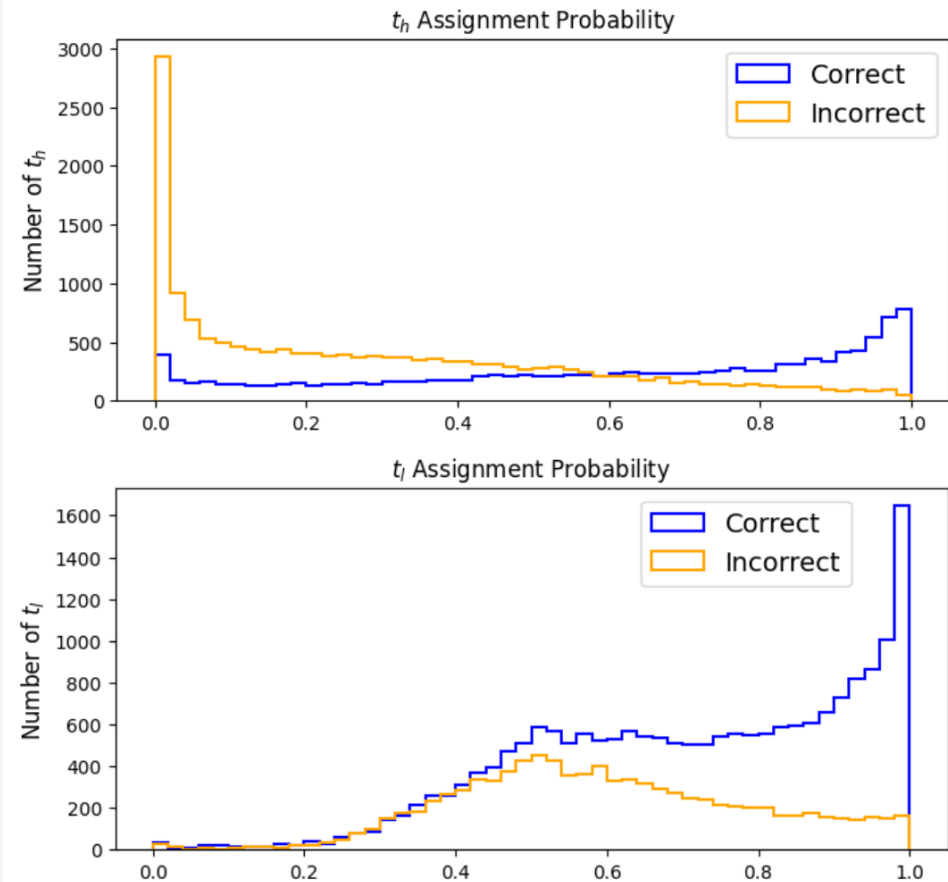

Event Type	N jet	Event fraction	Hadronic top: t_h	Leptonic top: t_l	Hadronic top (add vars)	Leptonic top (add vars)
All Event	==6	7.9%	0.326	0.752	0.332	0.766
	==7	19.8%	0.375	0.727	0.386	0.735
	>=8	72.9%	0.387	0.692	0.402	0.700
	inclusive	1	0.383	0.703	0.397	0.711

- **200% improvement compared to χ^2 method**
- Adding $\tau_1, \tau_2, \tau_3, \tau_4, \tau_5$ and ratio of charge particles over neutral particles (shown in brown column)
  - Small improvement (percent level)
- t_l efficiency drops to 0.632 if require both b-quark and lepton

# SPA-Net Assignment Probability

- The assignment probability can help to evaluate the output of the SPA-Net
  - Each top quark from SPA-Net prediction is classified as correct or incorrect by comparing it to truth label
  - Threshold on the assignment probability can help to improve the proportion of correct assignments

Result of HyPER

- Hypergraph can derived similar results in hadronic top quark reconstruction
- Parameter-efficient compared to SPA-Net: 500k vs.16M
- Performance reaches **~47%** in hadronic top reconstruction
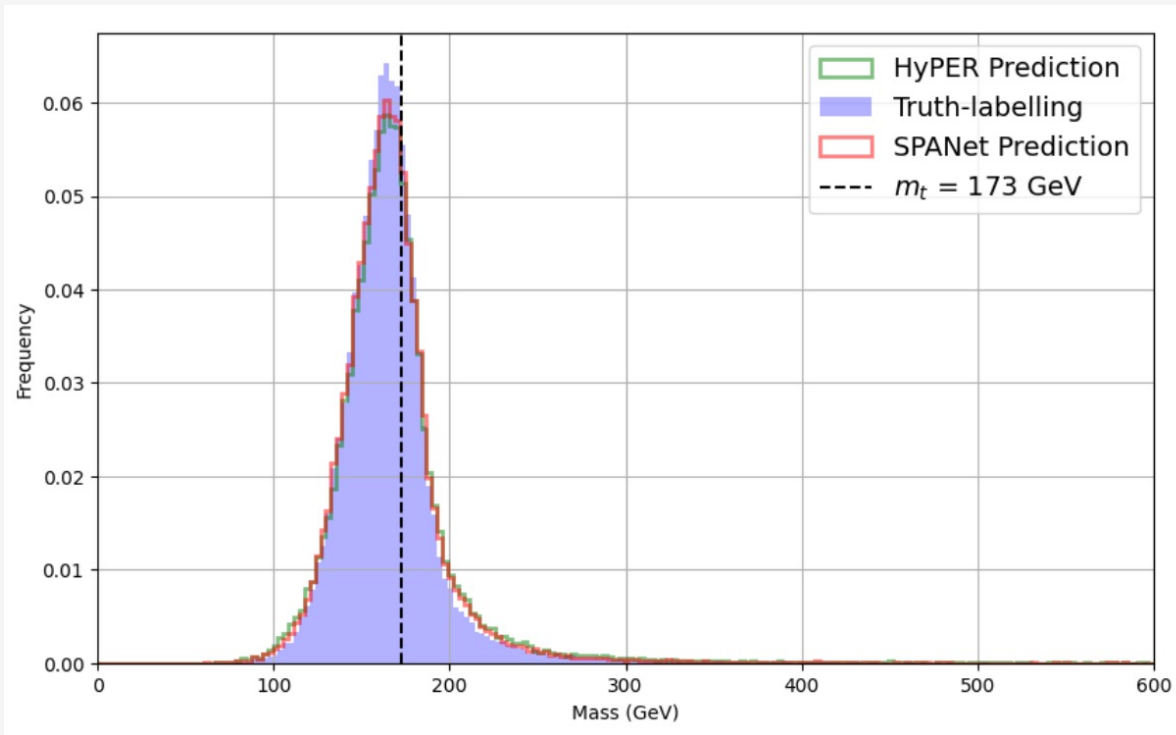
Event Type	N jet	SPANet $(2t_h+2t_l)$	SPANet $(2\ t_h)$	HyPER $(2\ t_h)$
All Event	==6	0.326	0.540	0.559
	==7	0.375	0.528	0.525
	>=8	0.387	0.461	0.453
	inclusive	0.383	0.471	0.467

- HyperGraph has no neutrino regression and can only reconstruct t_h
- SPA-Net and HyperGraph have similar performances when reconstructing t_h only
- Adding neutrino regression should help for HyperGraph

Multi-task Learning in Multi-lepton Channel

- In two lepton channel, the large missing ET coming from the neutrino arise the difficulty for reconstruction

- Machine learning method can regress or predict the momentum of the neutrino
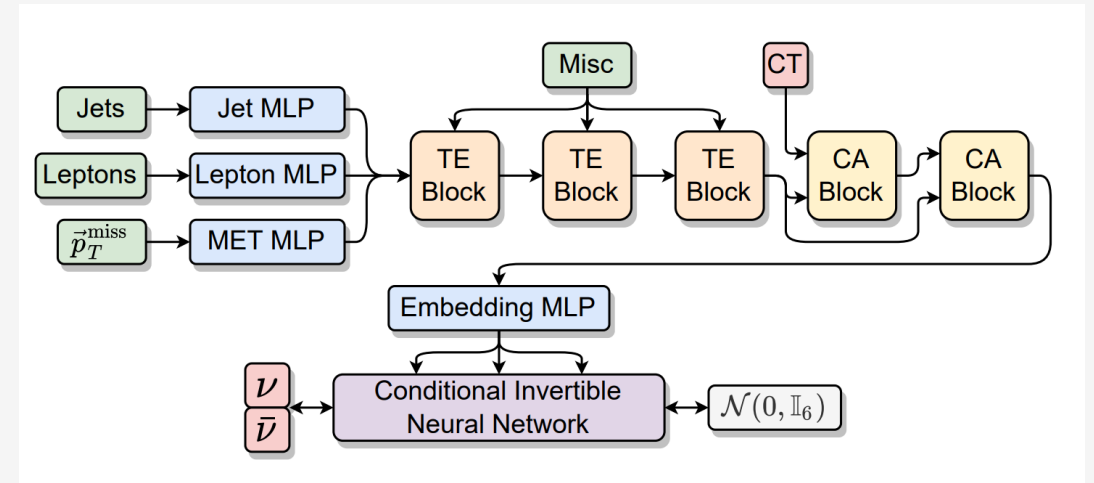
Commun Phys **7**, 139 (2024)   SciPost Phys. 12, 178 (2022)

PhysRevD.109.012005

- SPA-Net: use regression heads to output the value of the four momentum of neutrino
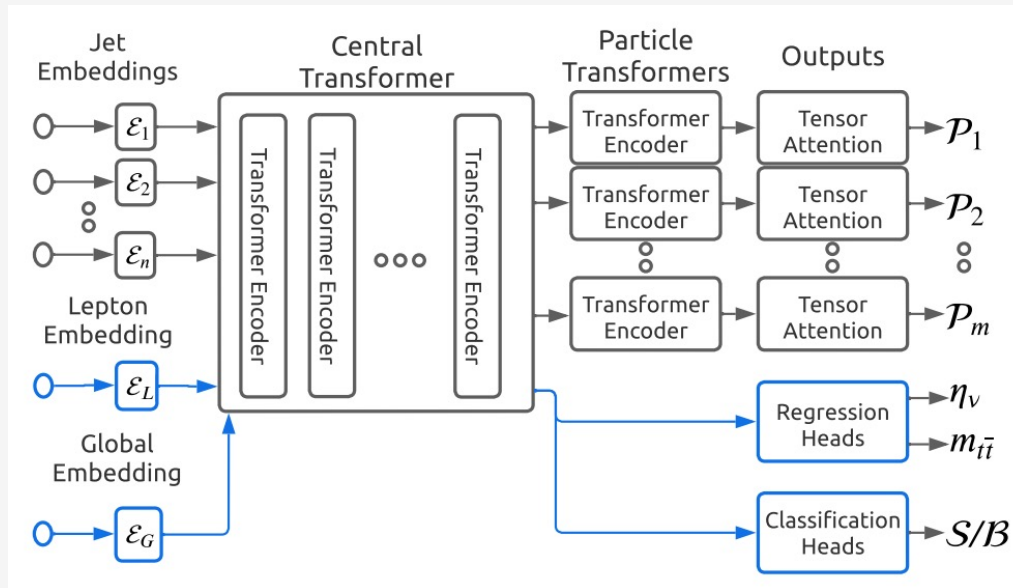
nu2Flow: use Normalization Flow to generate the distribution of the neutrino

SPA-Net Performance with Neutrino Reconstruction

- SPA-Net and Nu2Flow can model the neutrino's momentum well
- Using neutrino momentum information significantly improves top quark reconstruction efficiency
  - SPANet's neutrino regression improves t_h efficiency by 100% and t_l by 18%
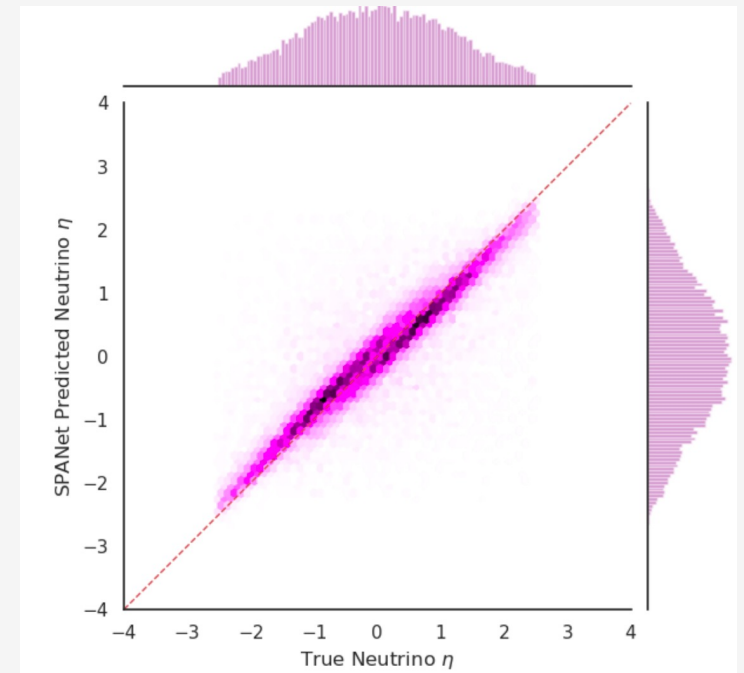
nu2Flow Neutrino Reconstruction
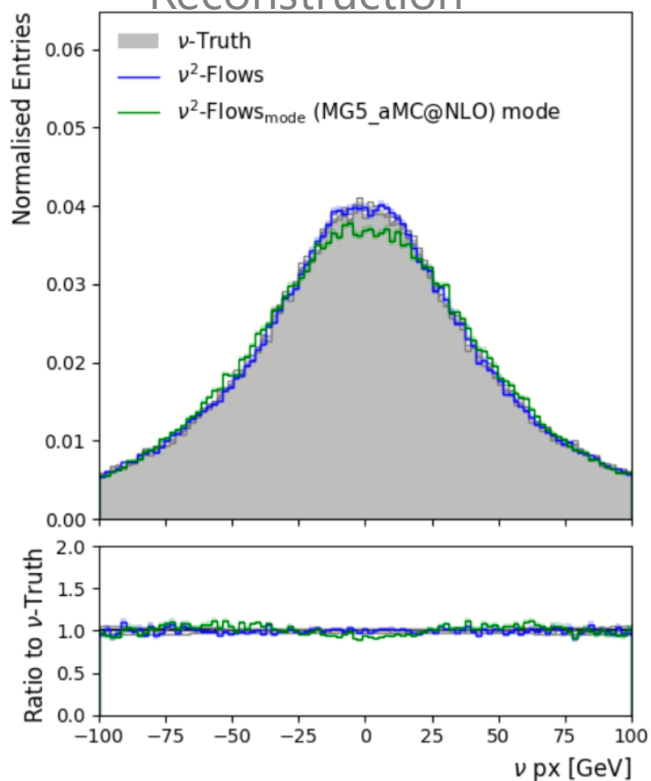
SPA-Net Neutrino Regression
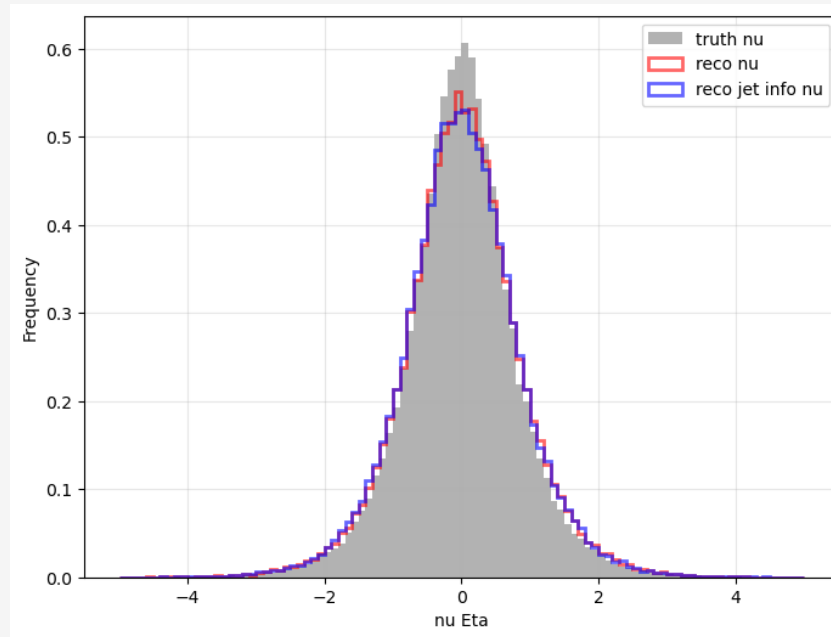
SPA-Net Neutrino Regression

Further Improvements to SPA-Net ?

- Lossing the b jet requirement in top quark reconstruction: ~20% improvements in hadronic top
- Adding neutrino momentum from nu2Flow as the input of the SPA-Net
  - Further improves t_h efficiency by ~10%
  - Expect smaller improvement in t_l efficiency (requiring both b-quark and lepton)

SPA-Net+neutrino regression

Event Type	N jet	Hadronic top	Leptonic top
All Event	inclusive	0.383	0.703

nu2Flow + SPA-Net

Event Type	N jet	Hadronic top	Leptonic top
All Event	inclusive	0.415	0.628

SPA-Net+no b-tagging requirement

Event Type	N jet	Hadronic top	Leptonic top
All Event	inclusive	0.502	0.701

Including large-R jets?
Better b-tagging?
Better MET/neutrino modeling?

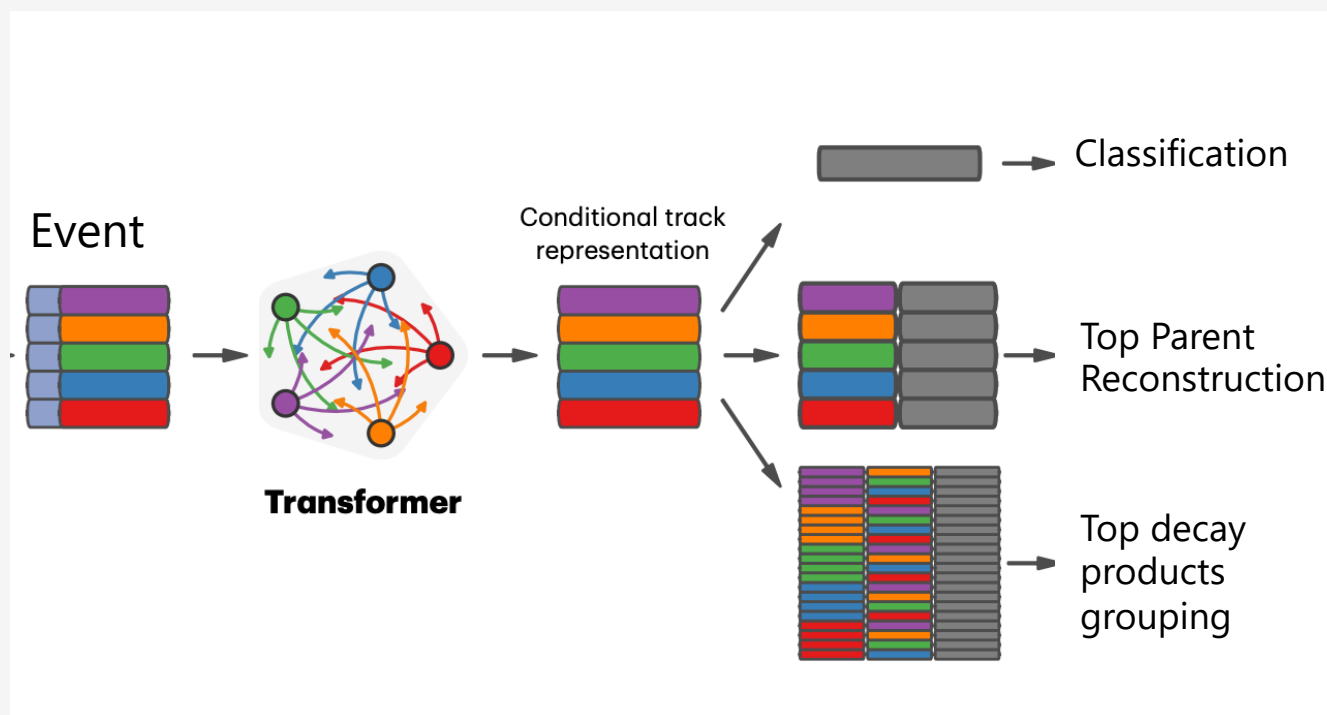# Four-top-quark Production in All-hadronic Channel

- Although SSML channel has better sensitivity due to less background but its decay branch ratio is 13%, 2 times smaller than the BR of all-hadronic channel
- CMS derived 2.7 sigma in the all-hadronic channel, using data-driven estimation for QCD background
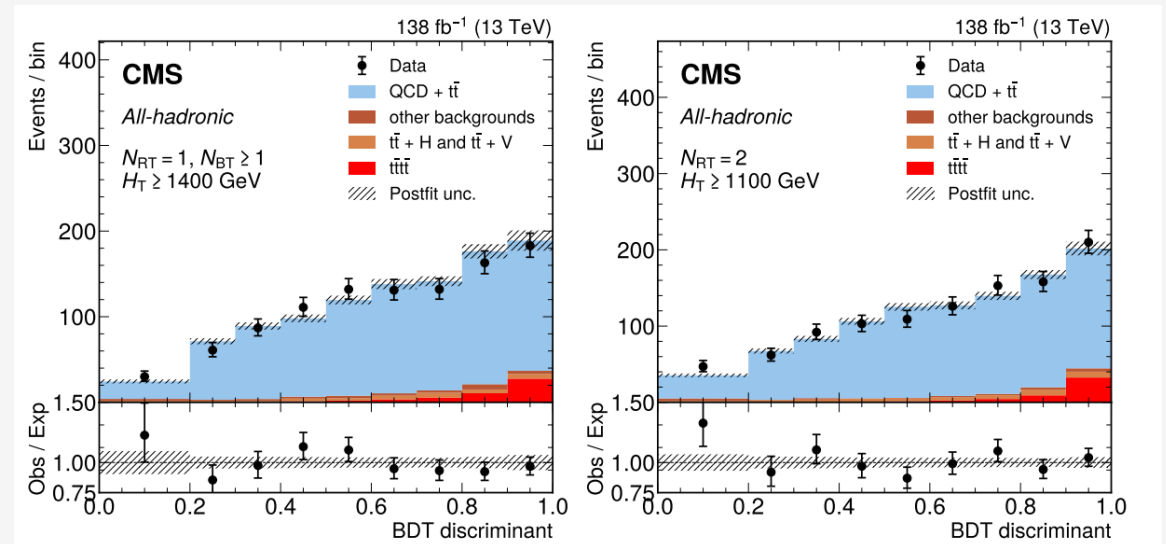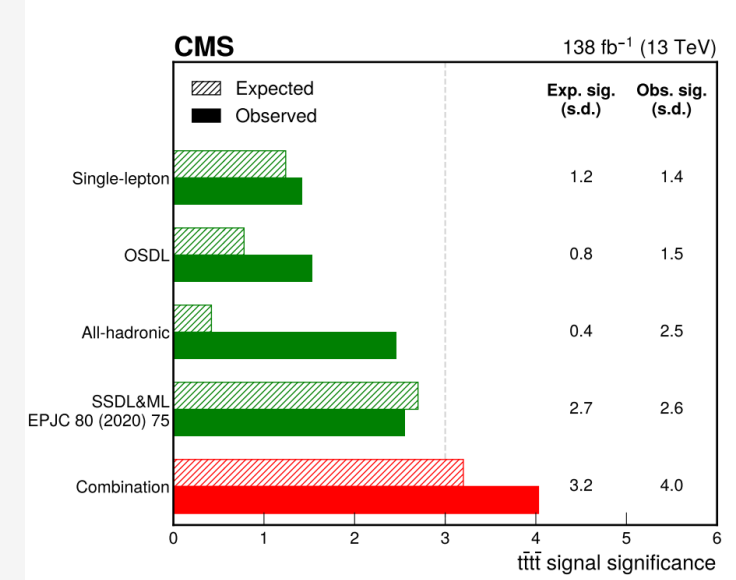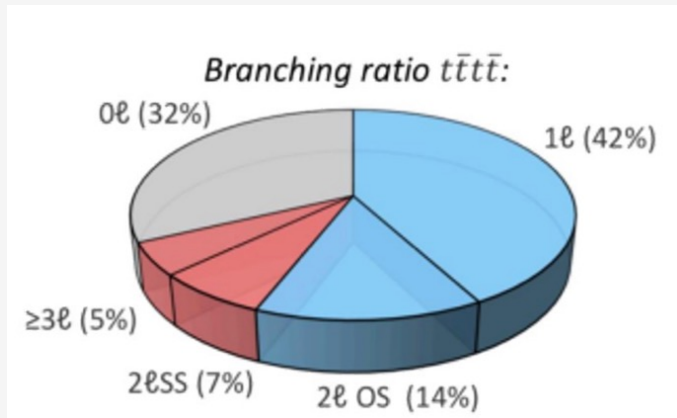
Transformer-based Framework for Four-top-quark Classification

Using the similar structure from the ATLAS transformer algorithm (GN2) but the input becomes the event level

- Considering dominated backgrounds, especially ttbar+jets
- Model inputs: jets kinematics and jet constituent-level information, event-level information

Four-top-quark Reconstruction in Full hadronic Channel

- Adding the top quark reconstruction task in the event-level GN2 framework
- Utilizing the top quark parent ID as the mask, trying to do the reconstruction like the image masking
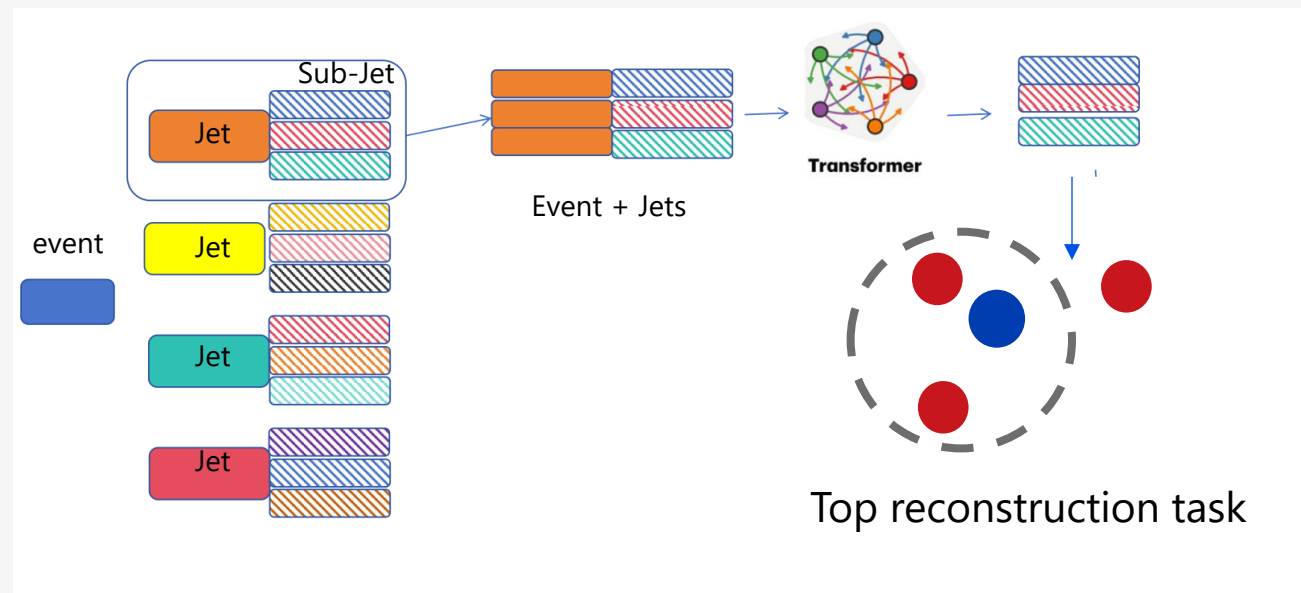
SPA-Net Algorithm

	N_{jets}	Event Fraction	SPA-NET Efficiency Event	SPA-NET Efficiency Top Quark
All Events	== 12	0.219	0.276	0.484
	== 13	0.304	0.247	0.474
	≥ 14	0.450	0.198	0.450
	Inclusive	**0.974**	**0.231**	**0.464**

SciPost Phys. 12, 178 (2022)

- SPA-Net reaches 46% reconstruction efficiency in all-hadronic channel
- Long training time due to large number of training parameters (16M)

GN2 Framework

Study on-going

Summary and Outlook

- **Top quark reconstruction can reduce the multi-top background in the analysis**

- **Machine learning method can be applied for the further top reconstruction in the different 4top final state**
  - Using SPA-Net for top reconstruction in SSML channel: **3 times** faster and hadronic top efficiency doubled to be **~40%,** leptonic top efficiency reaches **70%**
  - Further optimizations such as neutrino reconstruction and jet substructures

- **The sensitivity in the full-hadronic channel can be enhanced by the machine learning method**
  - Classification and top quark reconstruction task using the event-level GN2 algorithm

ATLAS
EXPERIMENT

Run: 304008
Event: 1533145462
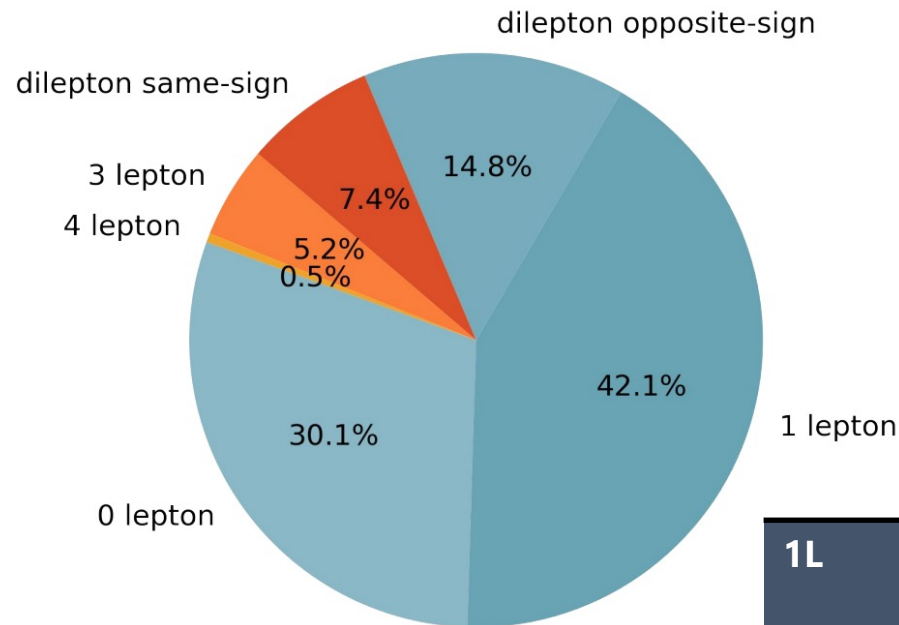2016-07-18 15:11:43 CEST
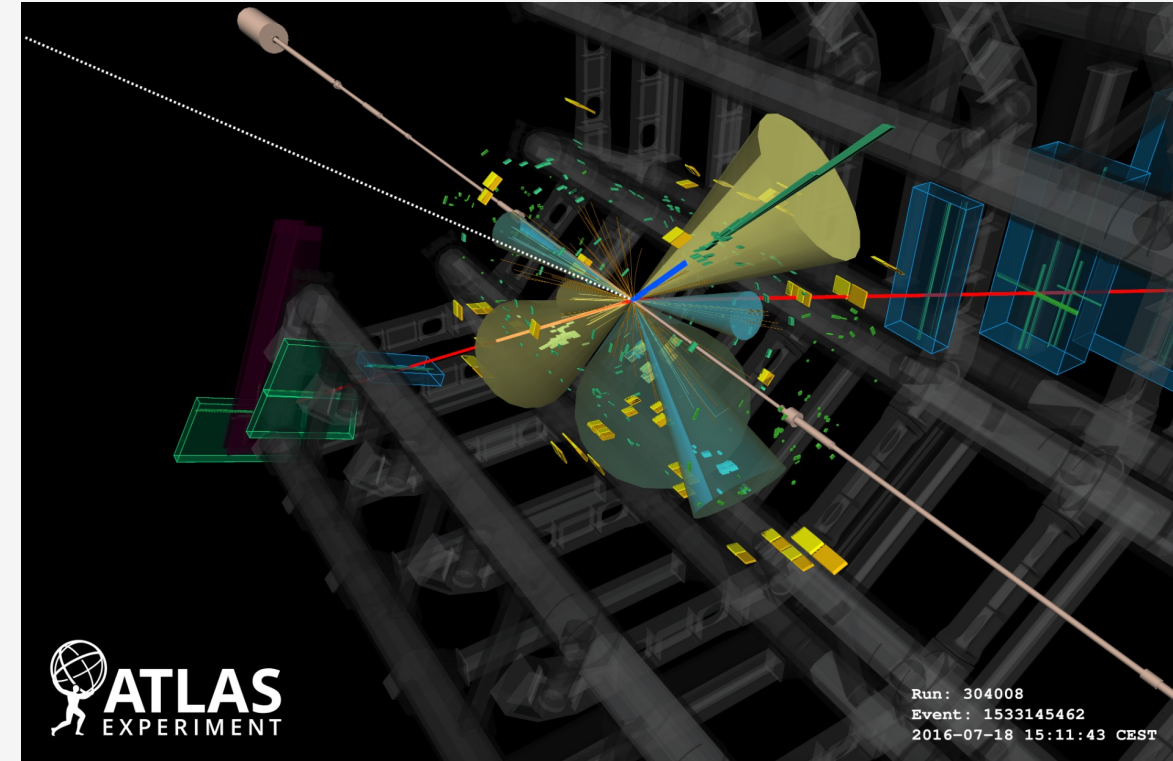
Backups

SPA-Net Performance in Different 4top channels

- Test the SPA-Net in different 4top channels
- 2L channel has the best performance overall in top quark reconstruction efficiency
- With the less hadronic top quarks, better performance

2L	Hadronic top	Leptonic top
Efficiency	0.383	0.703

All-had	Hadronic top
Efficiency	0.464

1L	Hadronic top	Leptonic top
Efficiency	0.305	0.604