# Top-antitop reconstruction

[a very biased overview]

:): 4 --x-- (Florence), 10/11/2023

**Baptiste Ravina**

ATLAS EXPERIMENT

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN IN PUBLICA COMMODA SEIT 1737

Alexander von Humboldt Stiftung / Foundation

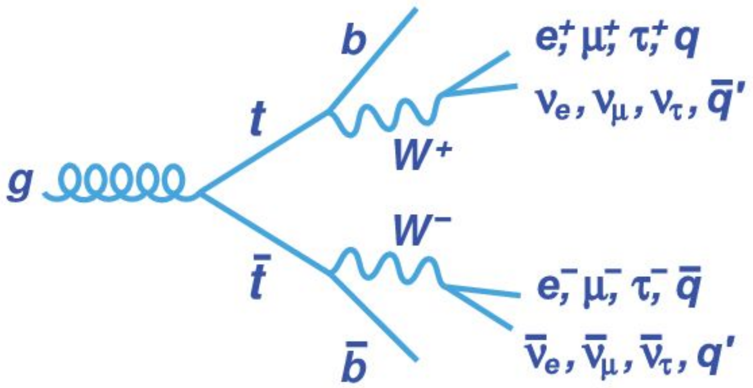
# For the next 20 minutes or so...

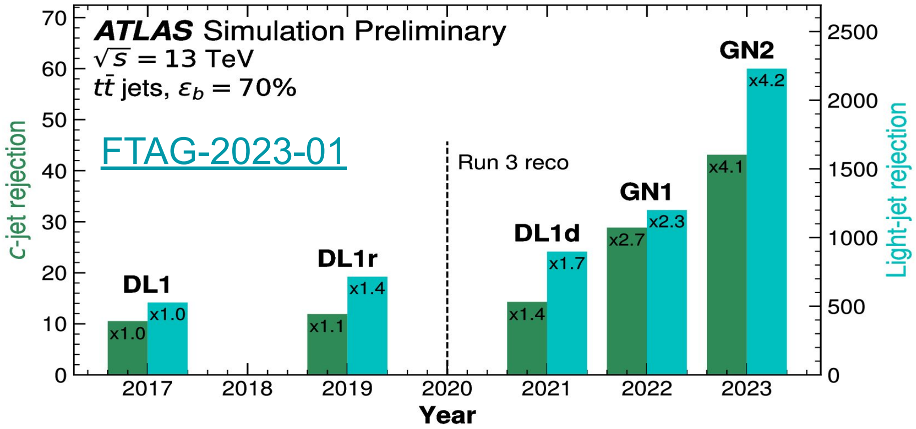


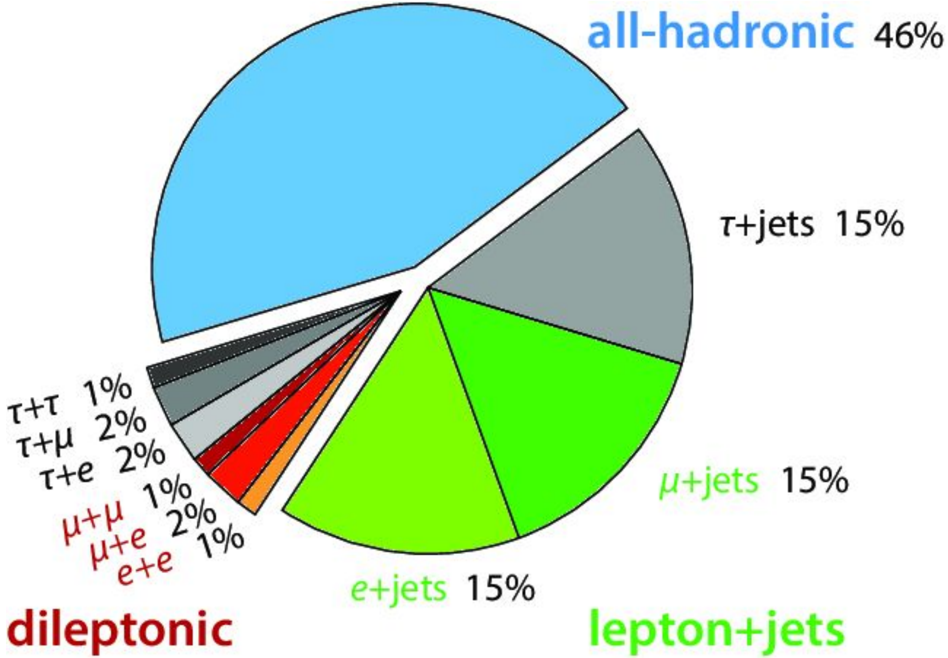
- I will discuss **some** approaches to top quark reconstruction
  - definitely not everything
- I will focus mostly on techniques used or soon-to-be-tried in **ATLAS**
  - hello, AJ
- I will not give you all the numbers, just an **overview** of what we can do
- I will talk disproportionately about **machine learning**
  - Jay won't be happy
- There is no real structure, just an unordered set of ideas
  - **Good luck.**

# The basics: what is top reconstruction?



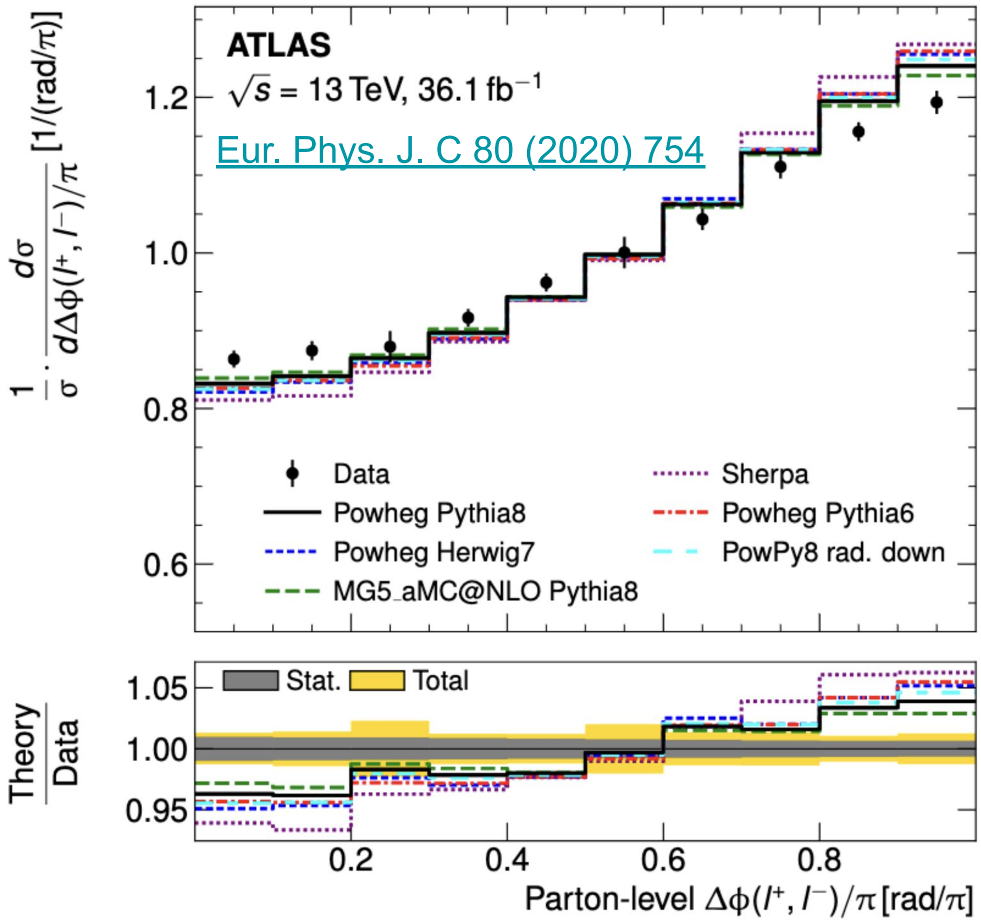


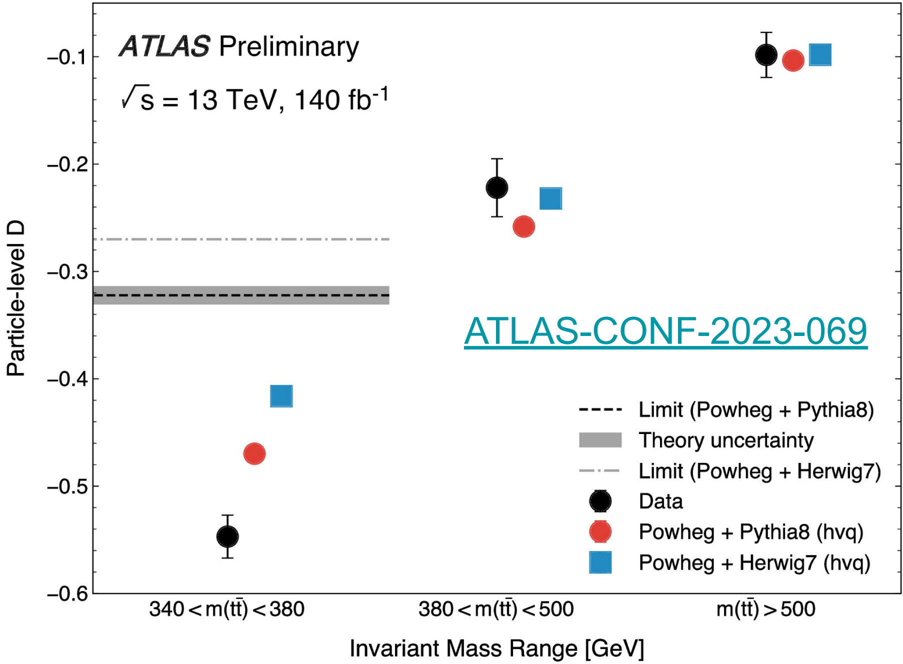



# The basics: why do we need top reconstruction?

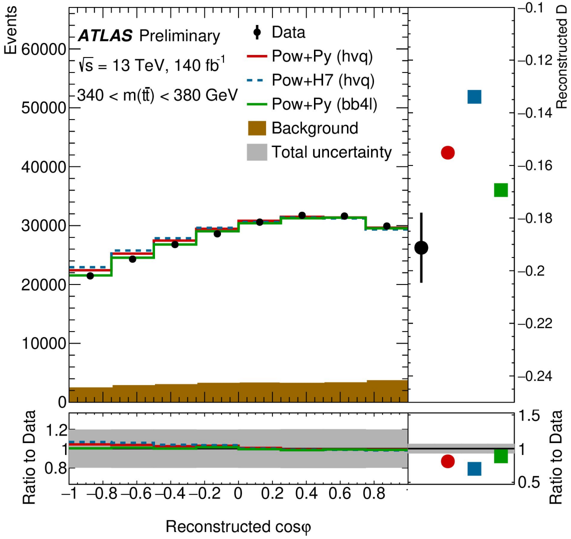




# Reconstruction for the dilepton entanglement result

the detector. Several methods are available to reconstruct the top quarks from the detector level charged leptons, jets and $E_T^{miss}$. The main method used in this work is the Ellipse method [70], which is a geometric approach to analytically calculate the neutrino momenta. Approximately 85% of events are successfully reconstructed by this method. If this method fails, the Neutrino Weighting method [71], which assigns a weight to each possible solution by the compatibility between the neutrino momenta and the $E_T^{miss}$ in the event, after scanning possible values of the pseudo-rapidities of the neutrinos, is used. If both methods fail,




ATLAS-CONF-2023-069

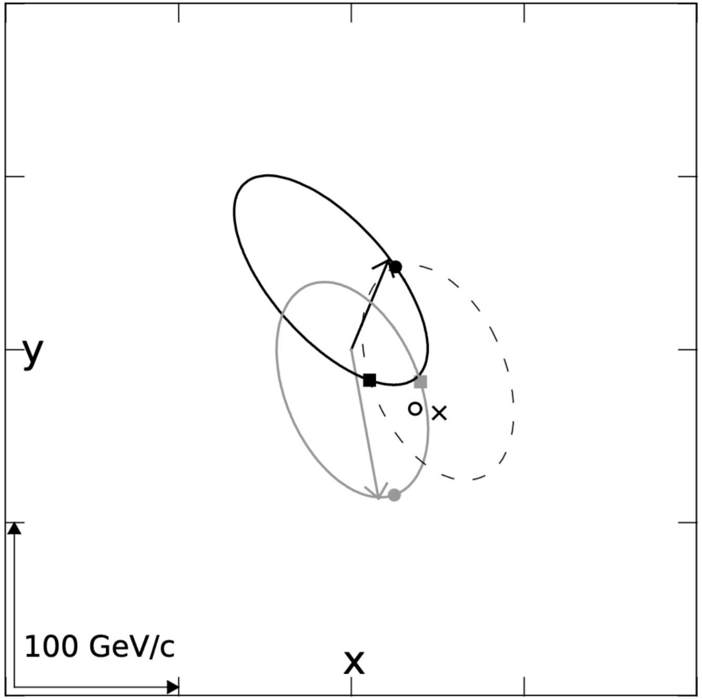
# The Ellipses method

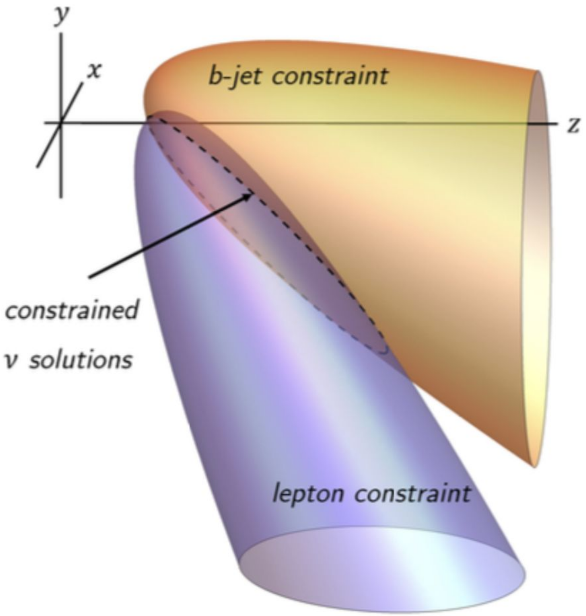
Nucl. Instrum. Meth. A 736 (2014) 169



**Assume:** everything is on-shell AND neutrinos are the source of the missing $E_T$

→ neutrino momenta are **geometrically** constrained to an ellipse



# The Neutrino Weighting method

Phys. Rev. Lett. 80 (1998) 2063



- Dates back to D0 (1997), they measured $m_{top} = 172.0 \pm 7.5$ GeV
- LHC Run 1 combination (2023) measured $m_{top} = 172.52 \pm 0.33$ GeV
- **Don't assume** that the missing $E_T$ comes from the neutrinos
  - instead **scan $(\eta_1,\eta_2)$** and for each pair extract $(p_{x1},p_{y1})$ and $(p_{x2},p_{y2})$ from the mass constraints
  - then compare to missing $E_T$ and extract a weight

$$w = \exp\left(\frac{-\Delta E_x^2}{2\sigma_x^2}\right) \cdot \exp\left(\frac{-\Delta E_y^2}{2\sigma_y^2}\right)$$

- Still have to check the b-jet assignments, possible dependence on $m_{top}$, smearing in case there are no solutions, …

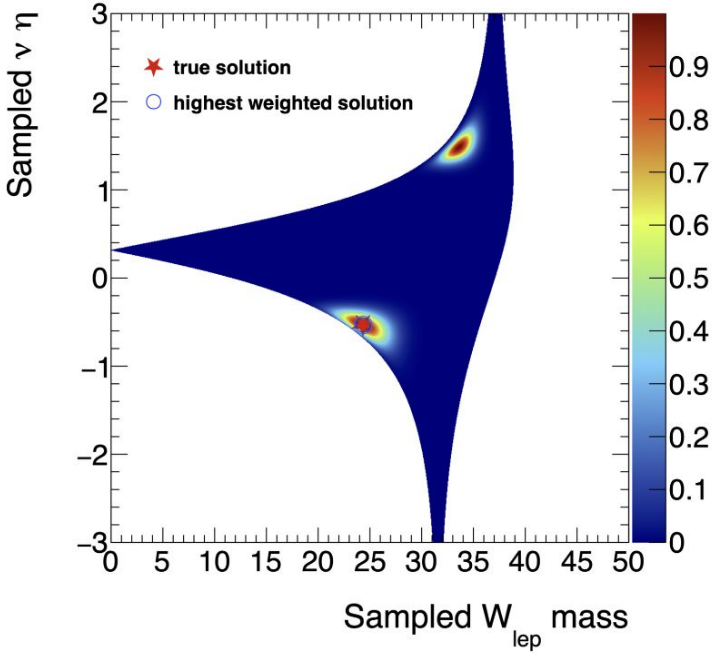
  → **very CPU-expensive!**

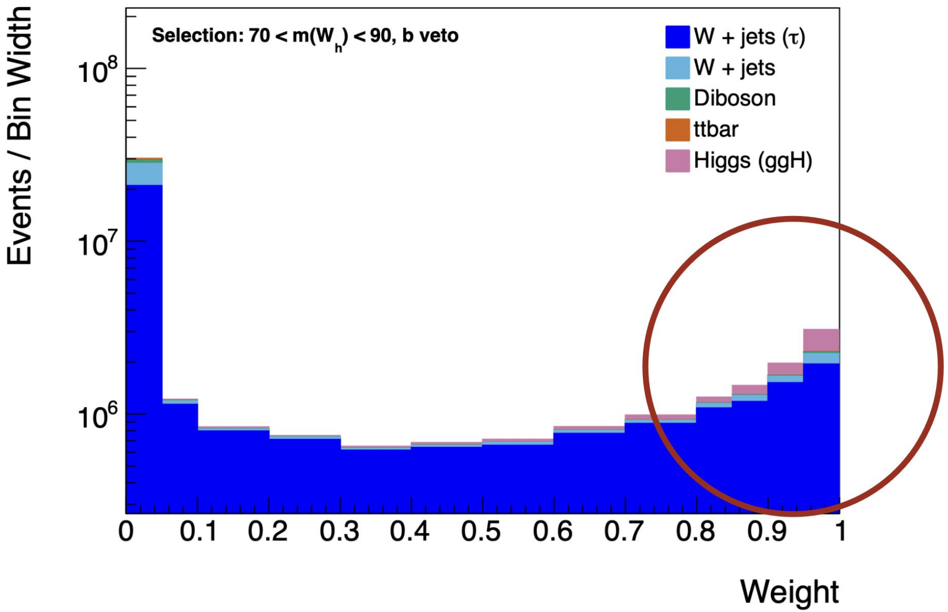
# Aside: Neutrino Weighter with a twist

See Theo Maurin's talk



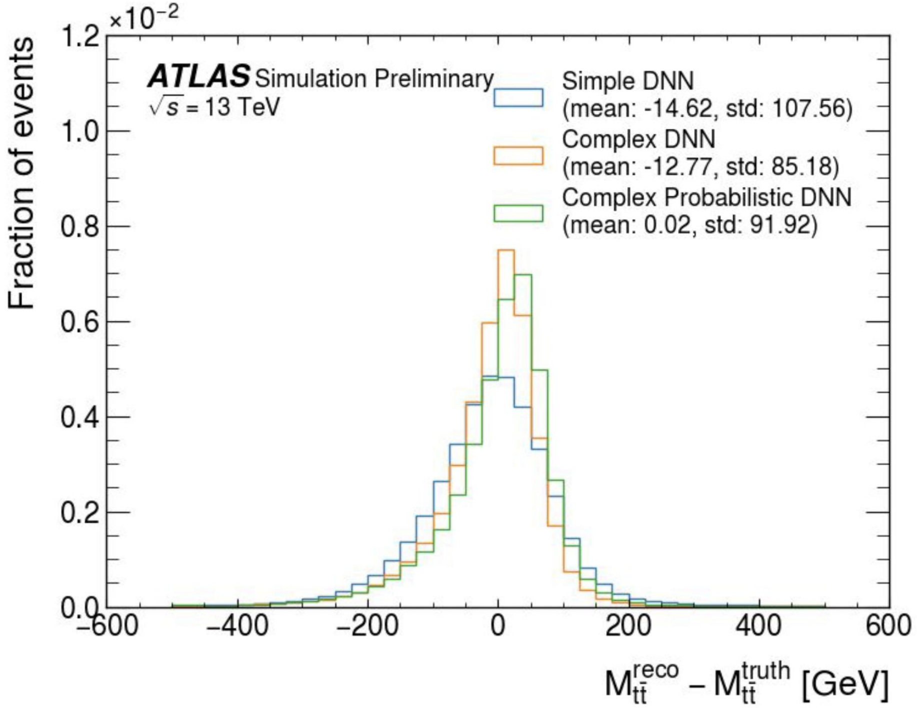
- **We reconstruct many Higgs each event under different assumptions of $m_{W^*}$ and $\eta_\nu$.**



# “Can we throw machine learning at it?”






**Simple** → **Complex**: add more inputs and more layers, get *improvement in resolution*.
**DNN** → **Probabilistic DNN**: get an estimate of the aleatoric uncertainty, *remove the bias*.

Reconstructing the two neutrinos’ 4-vectors is the hard part…

But maybe this is not always the goal?
For instance, we could regress m(ttbar) directly:

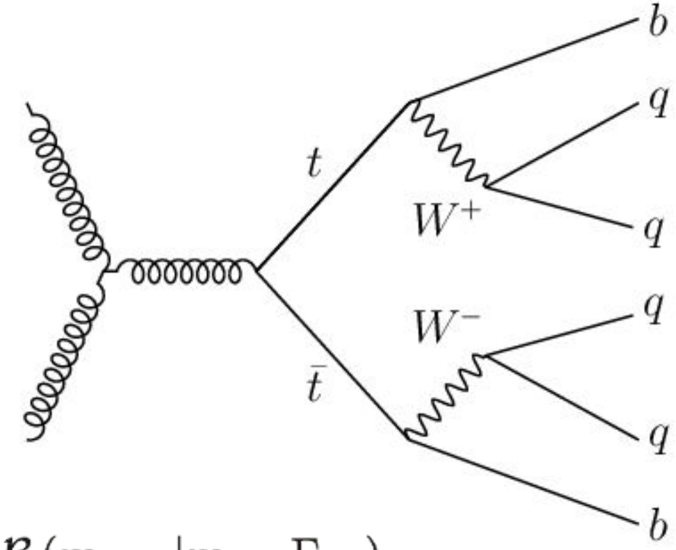
- Z’→ttbar resonance searches?
- dependence of m(ttbar) on top Yukawa?
- reducing the amount of dilution in QE/BIV measurements?

# All-hadronic ttbar: should be easy, right?



All decay products are visible jets → **completely avoid the problems** associated with neutrinos!

But now have to deal with combinatorics…




$$\chi^2 = \frac{(m_{b_1q_1q_2} - m_t)^2}{\sigma_t^2} + \frac{(m_{b_2q_3q_4} - m_t)^2}{\sigma_t^2} + \frac{(m_{q_1q_2} - m_W)^2}{\sigma_W^2} + \frac{(m_{q_3q_4} - m_W)^2}{\sigma_W^2},$$

$$\mathcal{L} = \mathcal{B}\left(m_{q_1q_2q_3}|m_t,\Gamma_t\right)\cdot\mathcal{B}\left(m_{q_1q_2}|m_W,\Gamma_W\right)\cdot \mathcal{B}\left(m_{q_4q_5q_6}|m_t,\Gamma_t\right)\cdot\mathcal{B}\left(m_{q_4q_5}|m_W,\Gamma_W\right)\cdot \prod_{i=1}^{6} W_{\mathrm{jet}}\left(E_{\mathrm{jet},i}^{\mathrm{meas}}|E_{\mathrm{jet},i}\right)$$

**Suffer from CPU cost of permutations**

Nucl. Instrum. Meth. A 748 (2014) 18

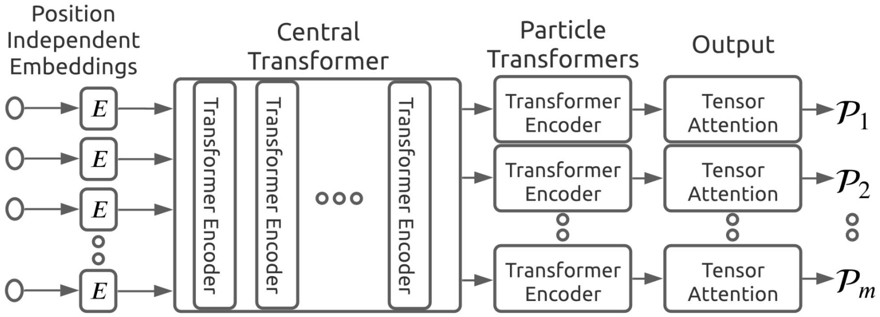
# Machine learning instead of combinatorics: SPA-Net



SciPost Phys. 12 (2022) 178

**S**ymmetry-**P**reserving **A**ttention **Net**work

**Transformer-Encoder:** state-of-the-art from Natural Language Processing
→ relate the input jets to each other in the latent space

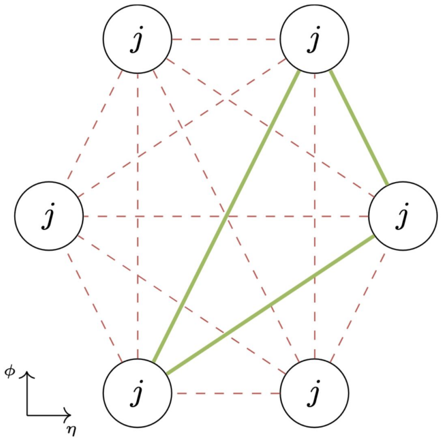


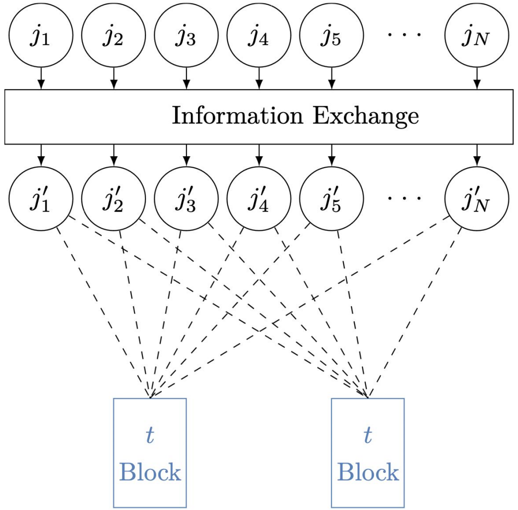

**Tensor attention:** impose symmetries of the topology
W ~ qq / top ~ bqq

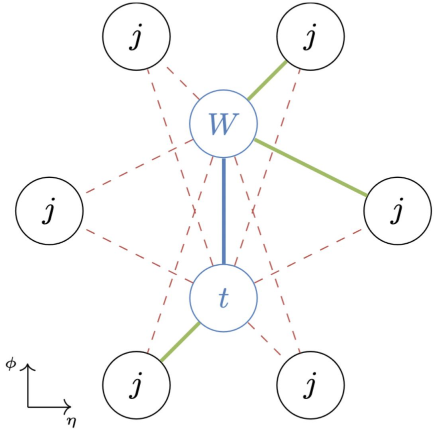
| | $N_{\text{jets}}$ | Event Fraction | SPA-NET Efficiency Event | SPA-NET Efficiency Top Quark | $\chi^2$ Efficiency Event | $\chi^2$ Efficiency Top Quark |
|---|---|---|---|---|---|---|
| All Events | == 6 | 0.245 | 0.643 | 0.696 | 0.424 | 0.484 |
| | == 7 | 0.282 | 0.601 | 0.667 | 0.389 | 0.460 |
| | $\geq 8$ | 0.320 | 0.528 | 0.613 | 0.309 | 0.384 |
| | **Inclusive** | **0.848** | **0.586** | **0.653** | **0.392** | **0.457** |
| Complete Events | == 6 | 0.074 | 0.803 | 0.837 | 0.593 | 0.643 |
| | == 7 | 0.105 | 0.667 | 0.754 | 0.413 | 0.530 |
| | $\geq 8$ | 0.145 | 0.521 | 0.662 | 0.253 | 0.410 |
| | **Inclusive** | **0.325** | **0.633** | **0.732** | **0.456** | **0.552** |

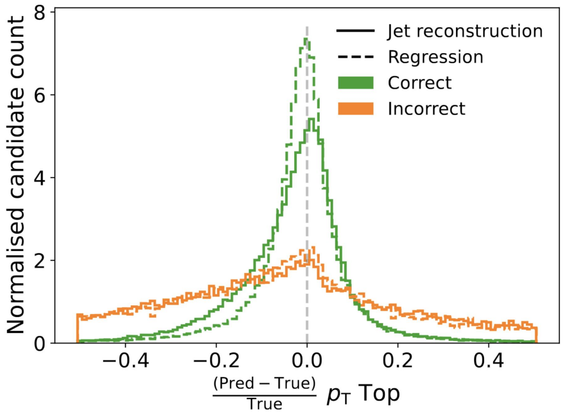
# Injecting yet more physics into the machine: Topographs

Phys. Rev. D 107 (2023) 11




**Physically motivated representation** of the inputs: **graph**
→ inject intermediate resonances and specify the allowed connections

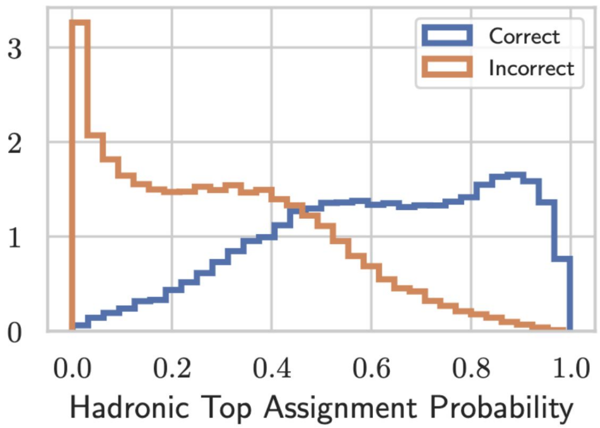
- Edge regression: find best assignments
- Node regression: predict the kinematics of the resonances
- Performs as well as SPA-Net

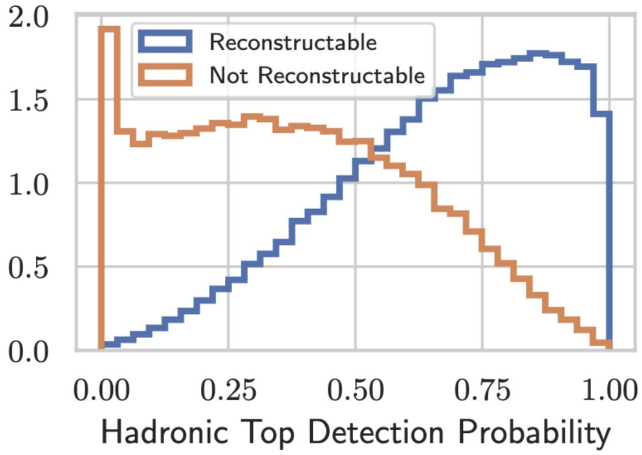
| | 6j 2b | 6j >=2b | 7j 2b | 7j >=2b | >=6j 2b | >=6j >=2b |
|---|---|---|---|---|---|---|
| Best Spanet [%] | 81.58 | 79.60 | 65.09 | 63.09 | 68.95 | 66.20 |
| Best Topograph [%] | 81.44 | 79.53 | 64.91 | 62.81 | 68.86 | 66.24 |

# From reconstruction to classification

arXiv:2309.01886

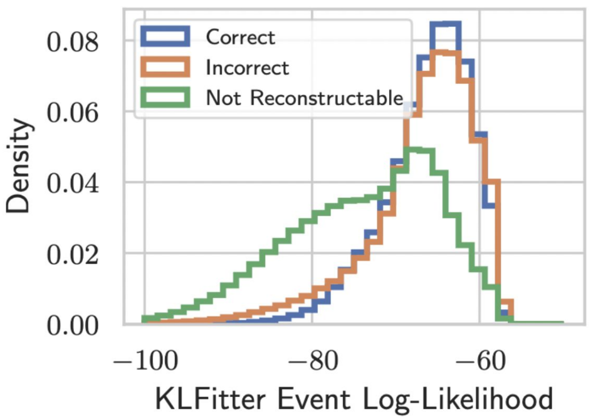


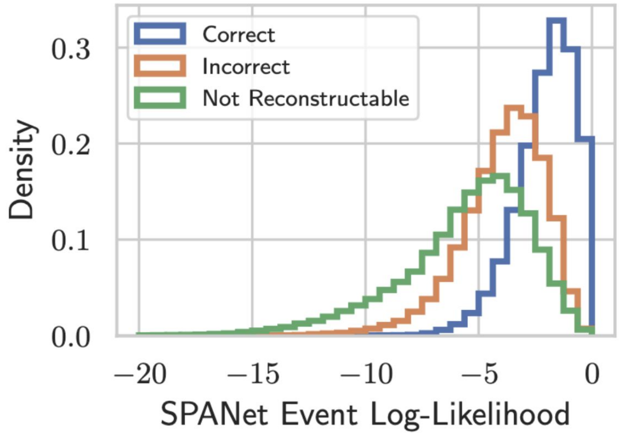


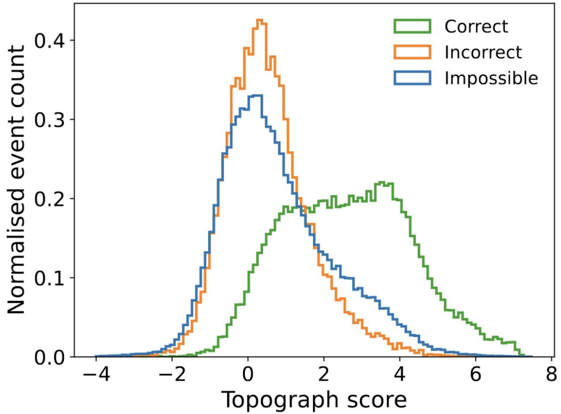




Could select only those events that are **well-reconstructed**:

- signal vs background?
- unfolding?
- modelling uncertainties?



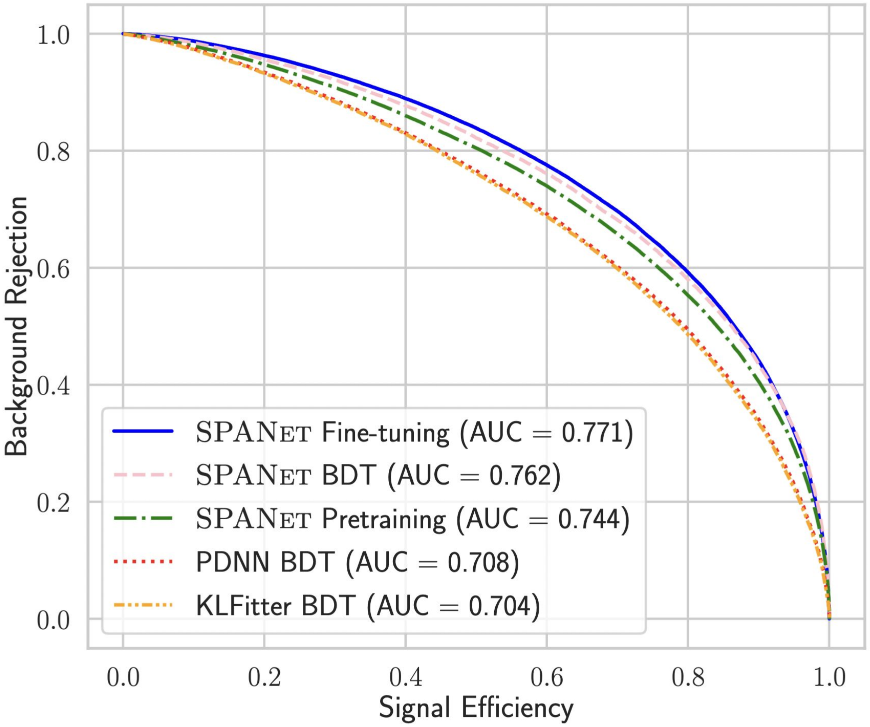



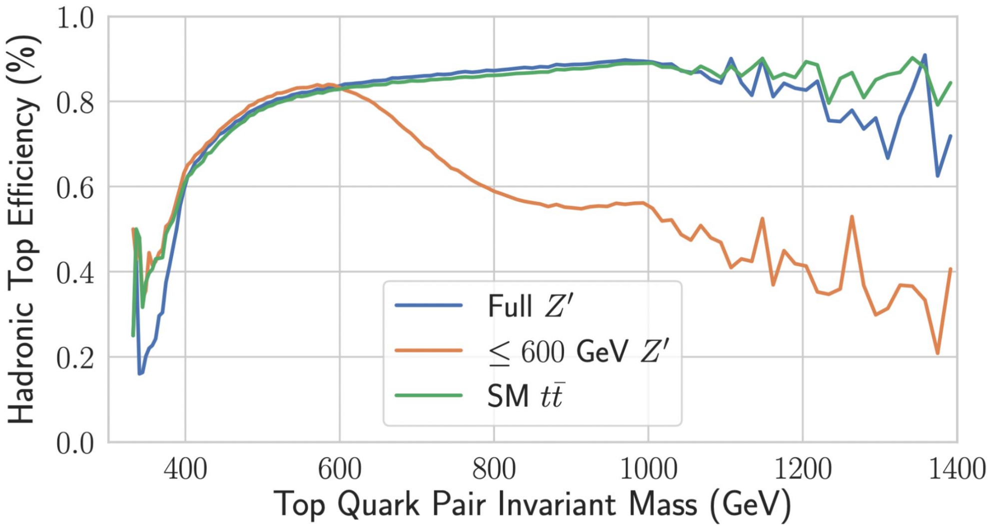



# Aside: other topologies

arXiv:2309.01886

### *ttH*(→*bb*) vs *tt*+*bb*




| | $t\bar{t}$ | $t\bar{t}H$ |
|---|---|---|
| KLFitter | 24 events per second | 2 events per second |
| PDNN CPU | 2626 events per second | 51 events per second |
| PDNN GPU | 3034 events per second | 101 events per second |
| SPANet CPU | 705 events per second | 852 events per second |
| SPANet GPU | 4407 events per second | 3534 events per second |

# A middle ground? ttbar → lepton+jets



Final state with a single neutrino: can be **fully determined from one mass constraint** (on-shell W) → analytical solution(s)

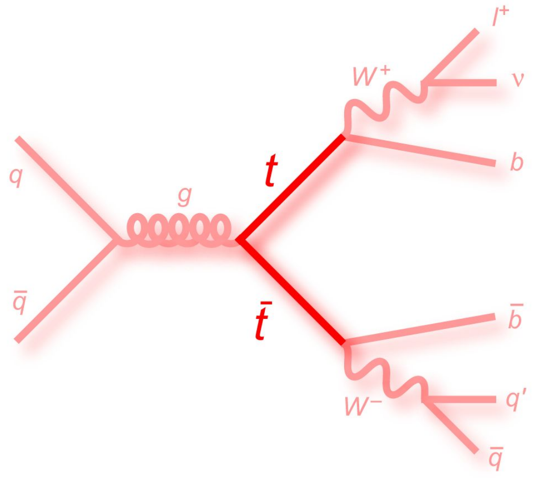
**Is this useful for spin correlation and quantum information studies?**

→ Yes! the d-quark from the W decay has $\alpha_{spin}$~1

- As seen in Theo Maurin's talk: can be accessed by c-tagging the other W-jet.
- Can also consider the "optimal hadronic direction" (Dorival Gonçalves's talk)

$$p_z^\nu = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a},$$

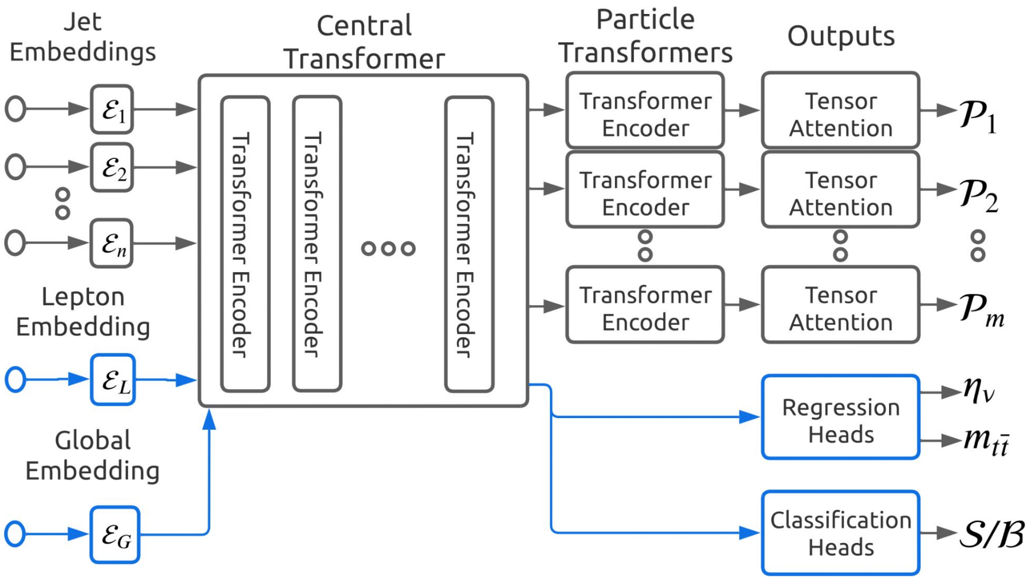
$$a = (p_z^\ell)^2 - (E^\ell)^2,$$
$$b = \alpha p_z^\ell,$$
$$c = \frac{\alpha^2}{4} - (E^\ell)^2 (p_\mathrm{T}^\nu)^2,$$
$$\alpha = m_W^2 - m_\ell^2 + 2(p_x^\ell p_x^\nu + p_y^\ell p_y^\nu).$$

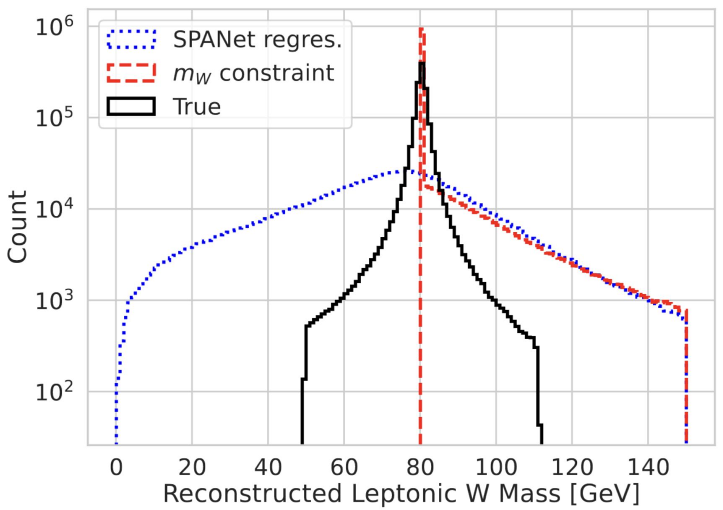


# SPA-Net with neutrinos

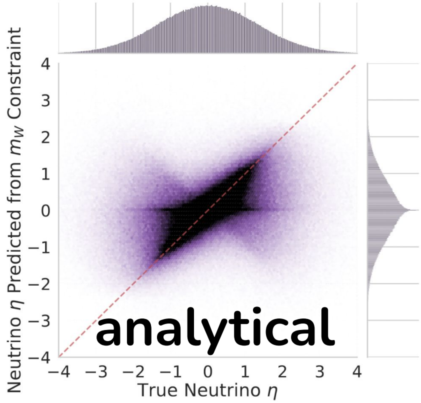
arXiv:2309.01886

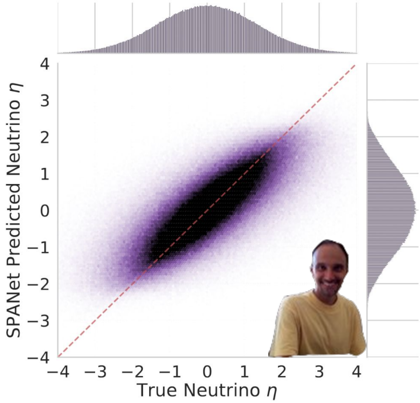




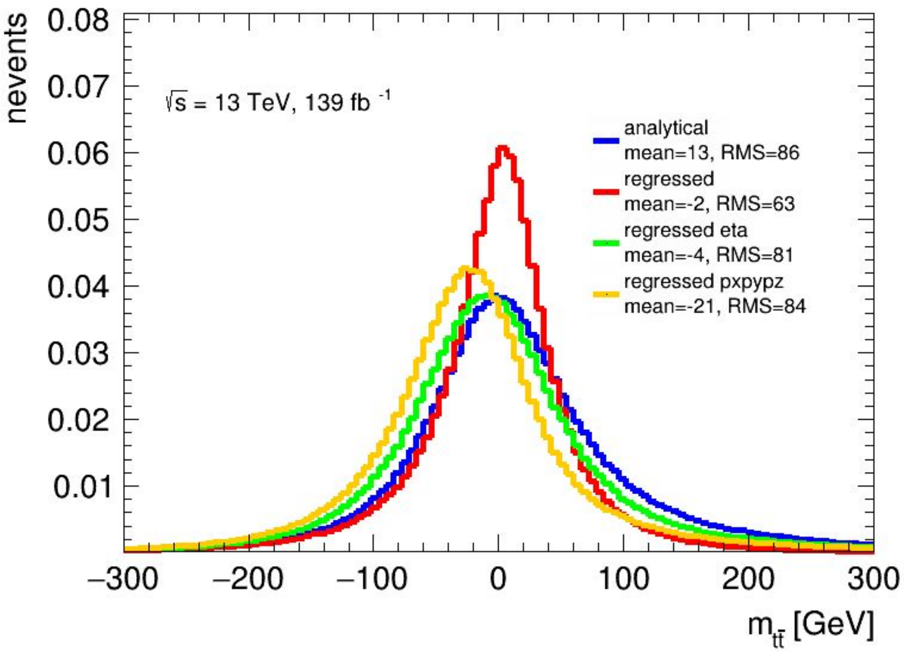

- Extend the jet-only model with specialised embeddings for **leptons** and **missing $E_T$**
- Output targets are now (bqq) and (b)
- Also add **regression** tasks

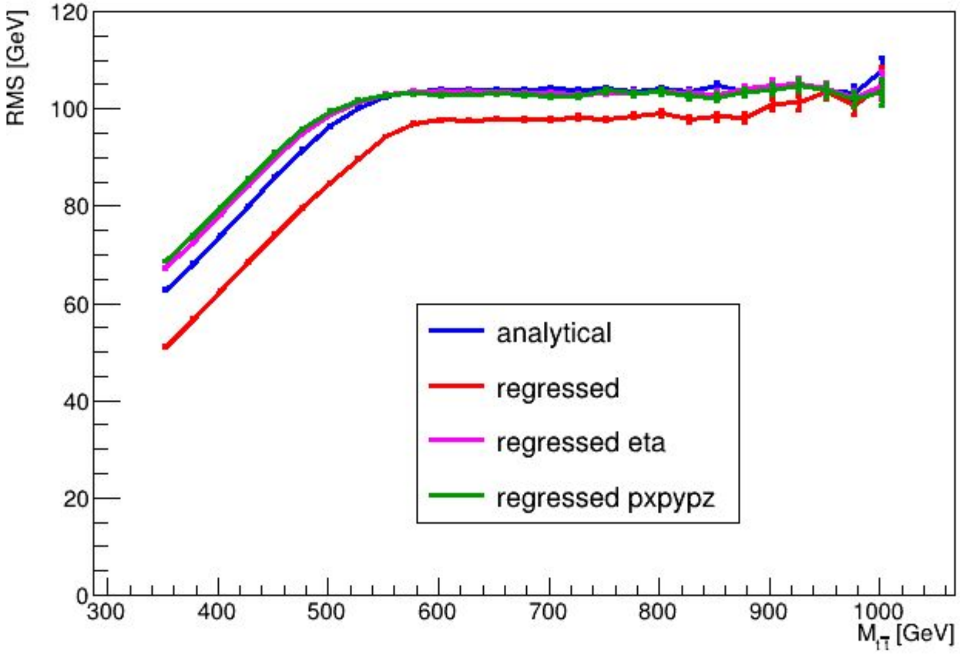


# SPA-Net with neutrinos: m(ttbar)





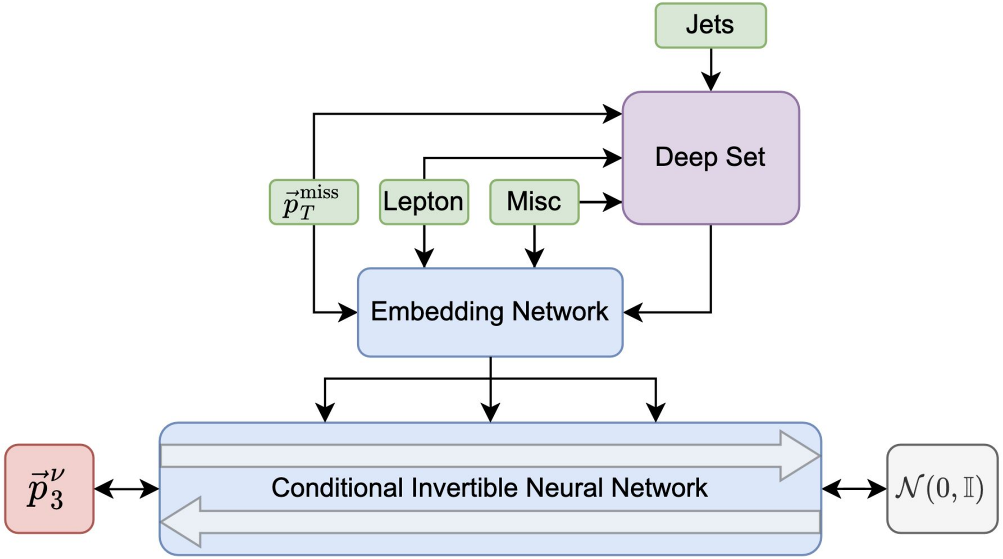

**"Simple guess"** of neutrino kinematics is not very useful
→ but maybe regression of m(ttbar) can help **select events for QE/BIV?**

# Conditional neutrino regression: $\nu$-flows

[SciPost Phys. 14 (2023) 159](https://scipost.org/SciPostPhys.14.6.159)



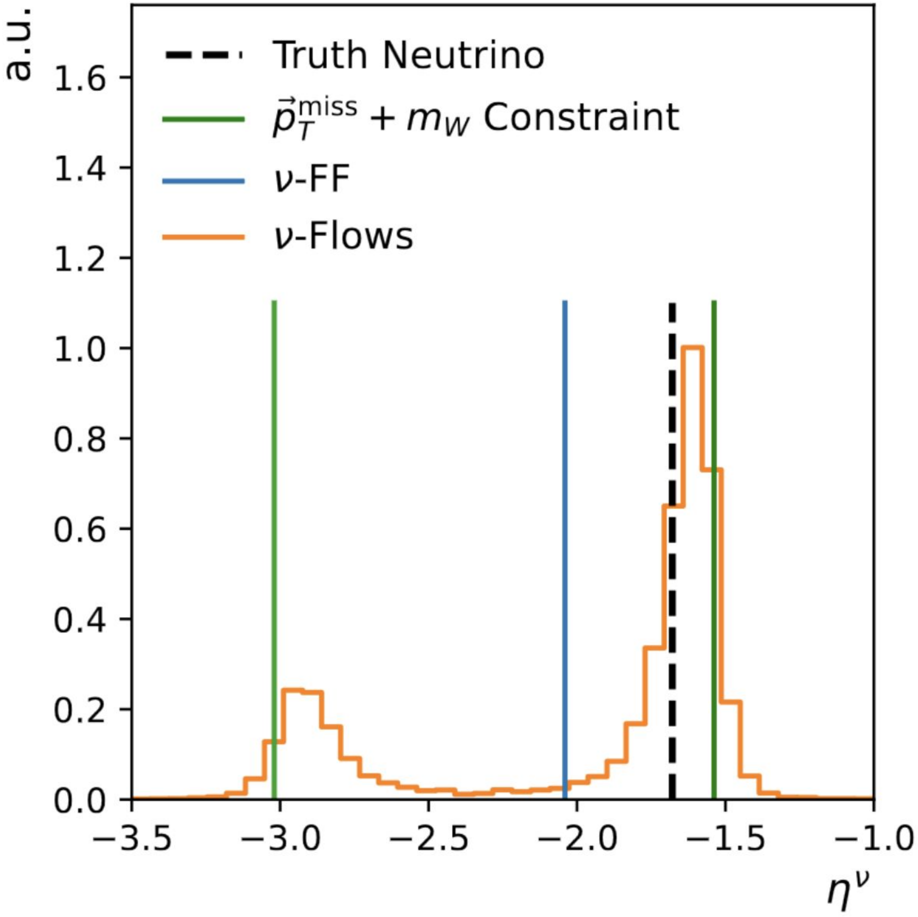
1. Embed your input particles in some way
2. **Train a mapping of the Normal distribution to the kinematics of the neutrinos**
3. Learn what the likelihood of the neutrino kinematics based on the rest of the event
→ no assumption of on-shell W's, perfect reconstruction etc.

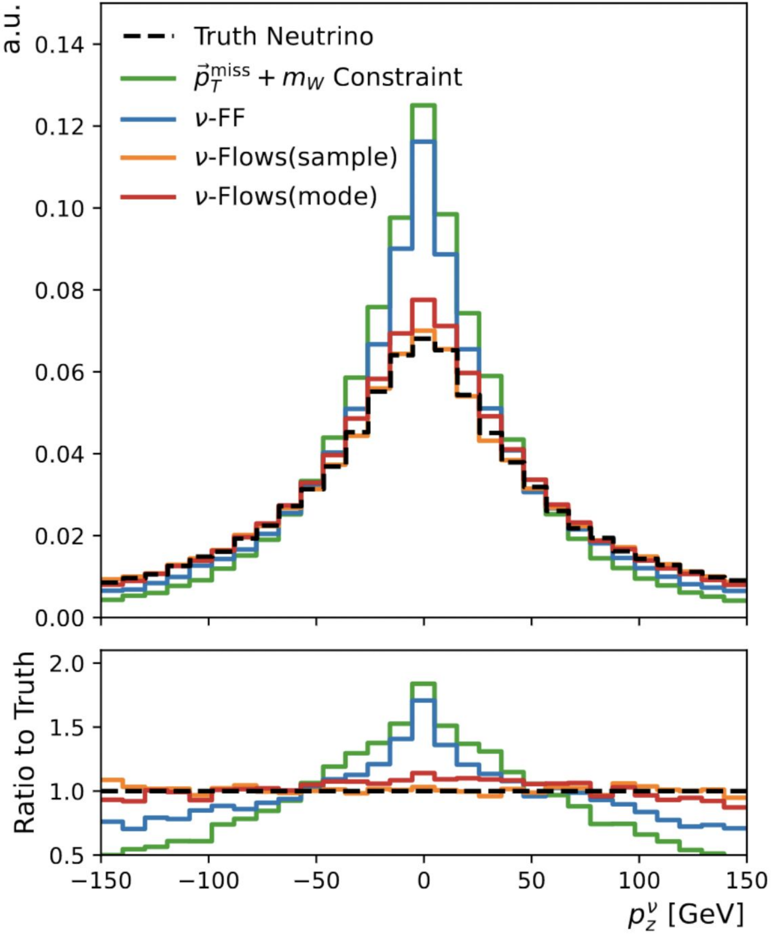


# Conditional neutrino regression: $\nu$-flows

SciPost Phys. 14 (2023) 159





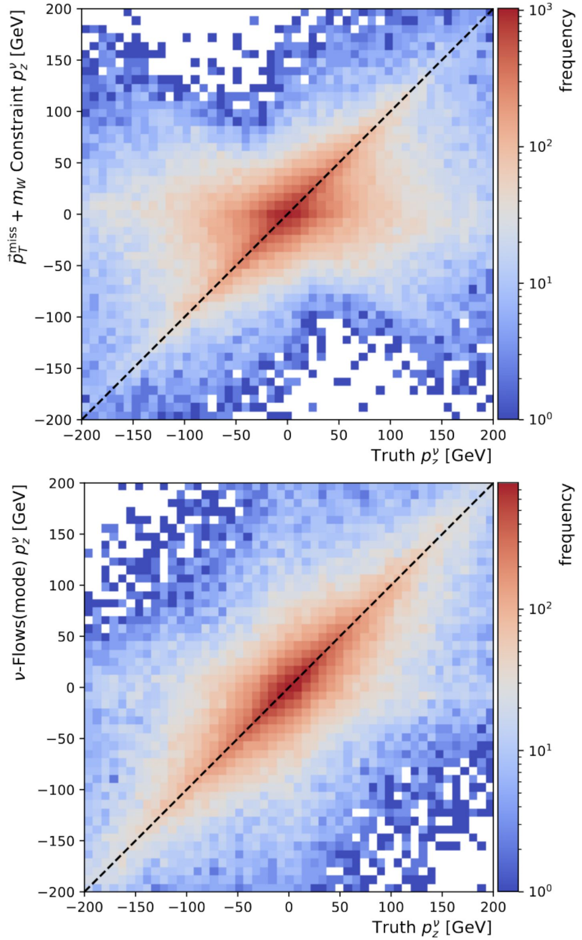
# Conditional neutrino regression: $\nu$-flows

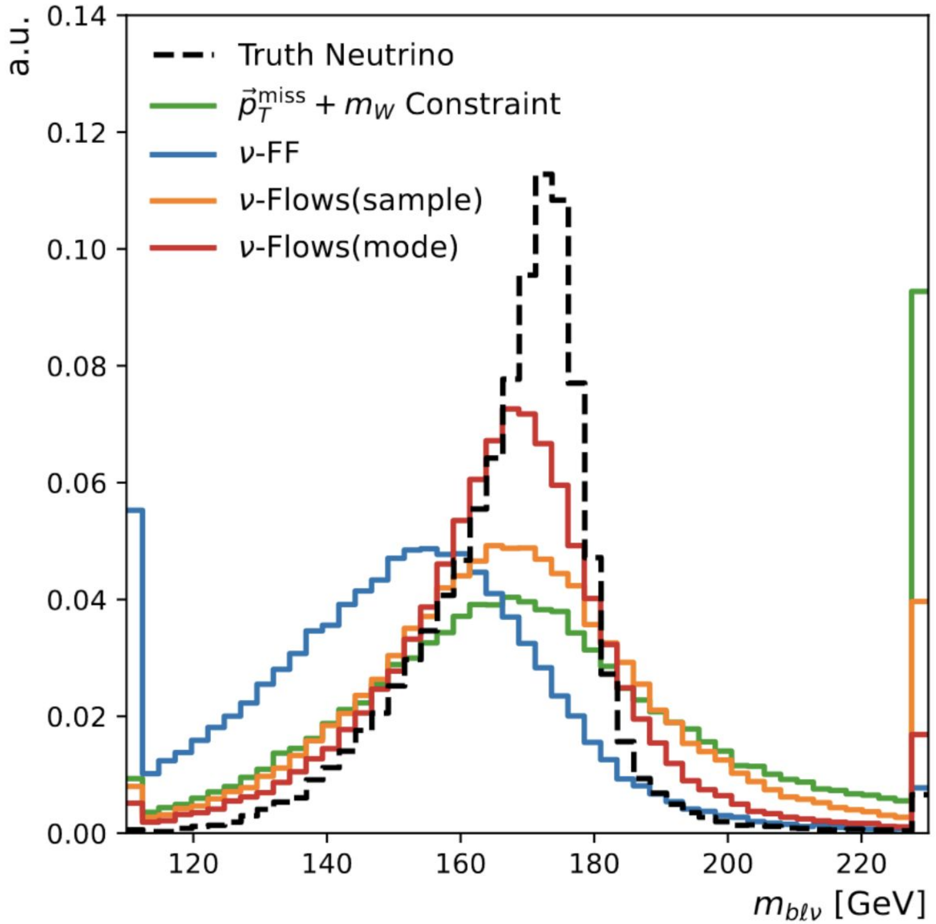
SciPost Phys. 14 (2023) 159



analytical

$\nu$-flows

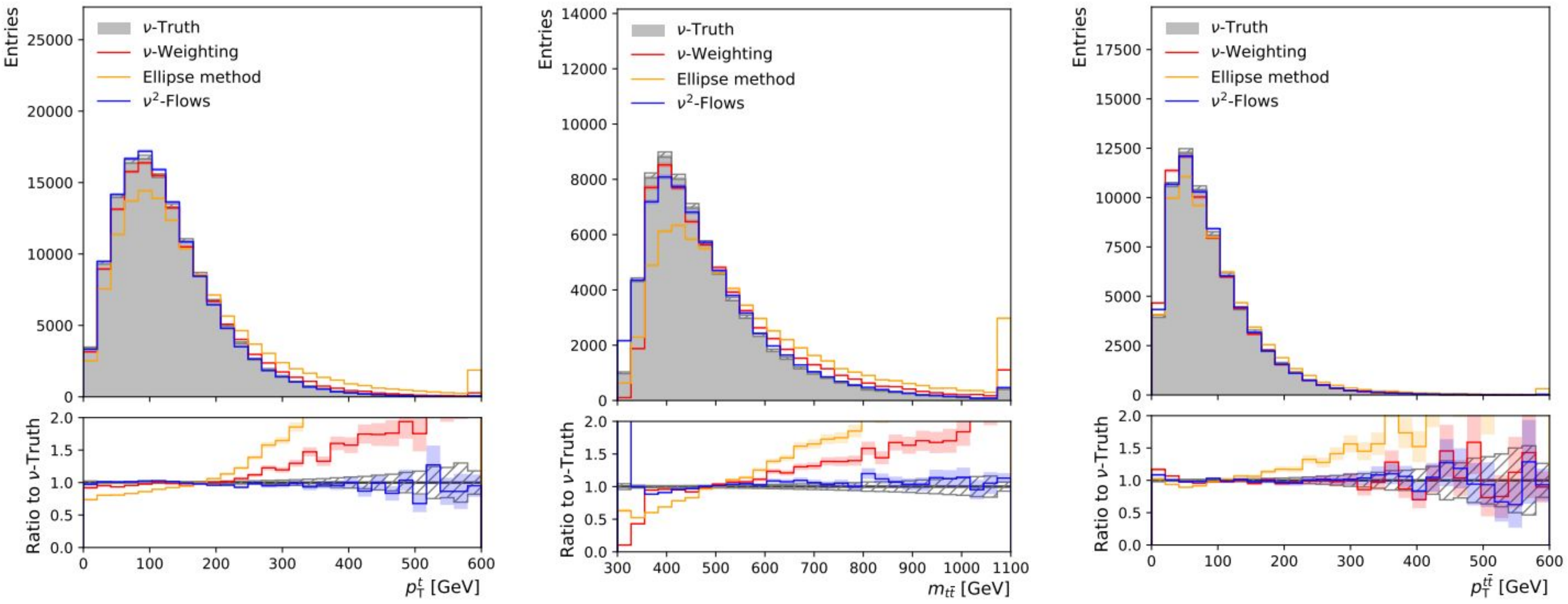


# More neutrinos! $\nu^2$-flows

arXiv:2307.02405

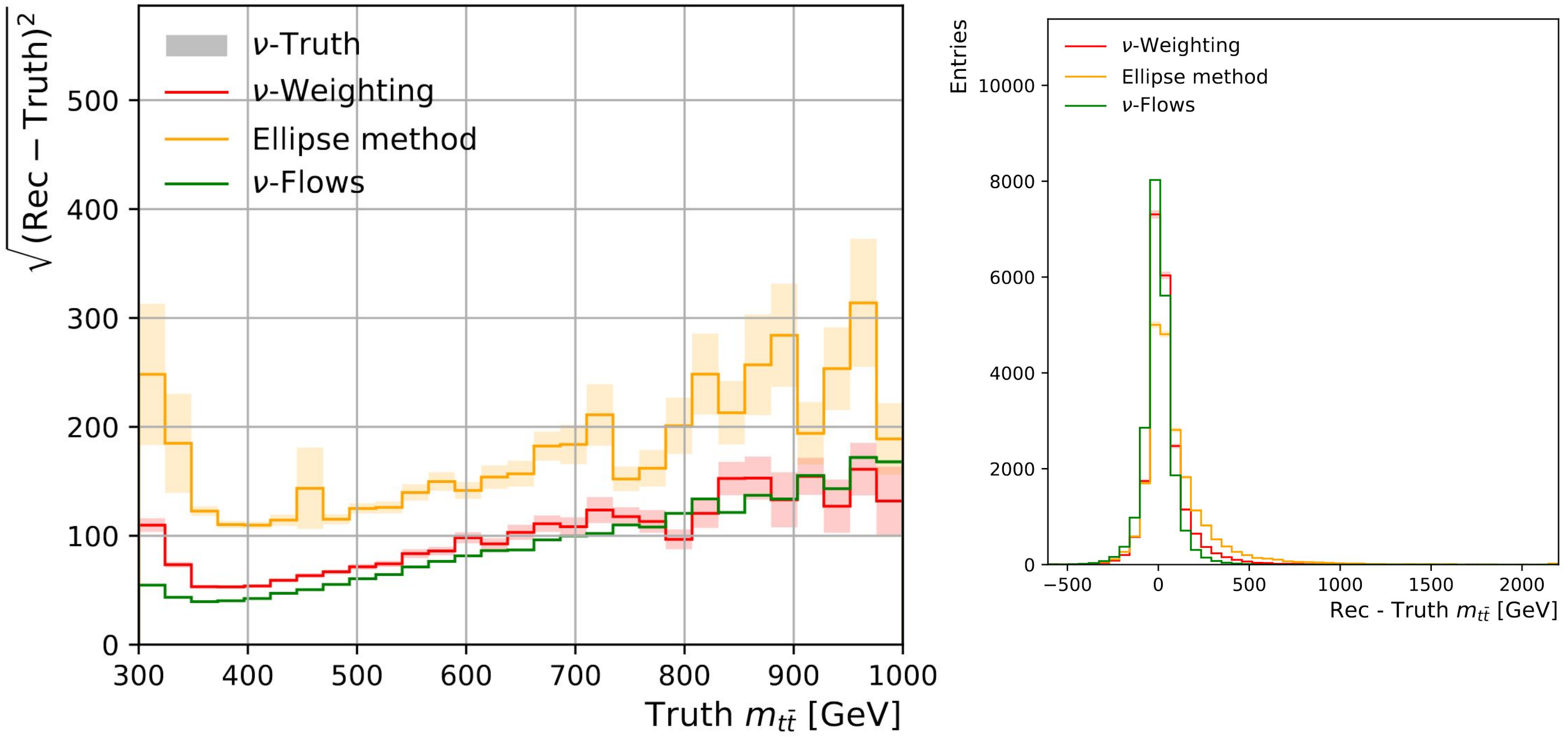




# More neutrinos! $\nu^2$-flows

arXiv:2307.02405

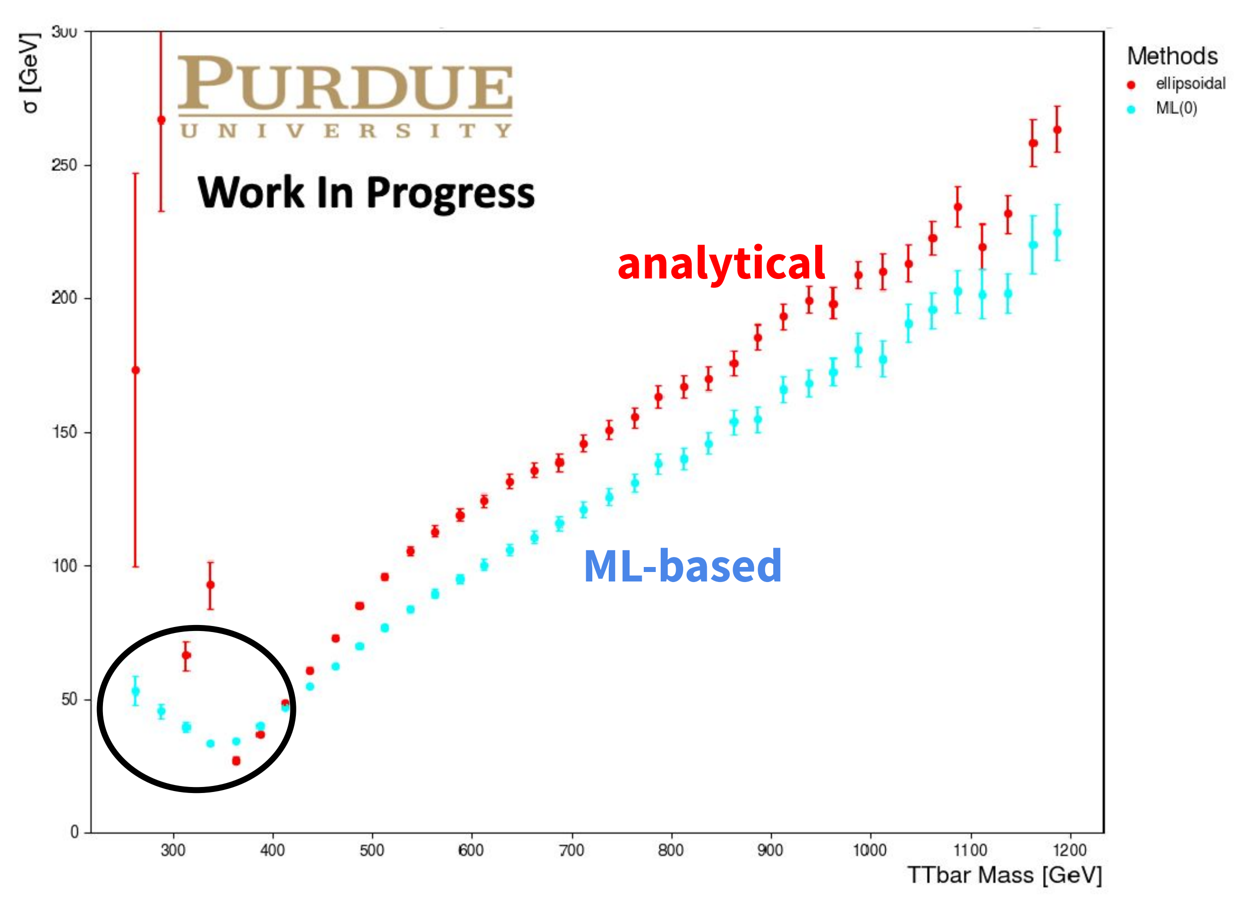




# What is CMS up to, these days?..

AJ Wildridge's talk at the Oxford workshop





# In conclusion...



- Didn't talk about ML for boosted tops / jet substructure
- **Reconstruction for classification** will be a very powerful tool
- Some algorithms target specific observables (e.g. m(ttbar))
- Others offer to **perform the full reconstruction**
  - harder to check for biases
  - do we trust the kinematic correlations in the absence of strong physics assumptions? (i.e. does ML know about 4-vectors)
- **Systematics** on top reconstruction: better or worse with ML?