Unravelling the origins of top pairs: production mode analysis in ATLAS

DESY Summer School Program project

Valerio Tinari
ATLAS group

Supervised by: *K. Behr, E. Jones, M. Gonzalez , F. Jolly*

02/09/2025

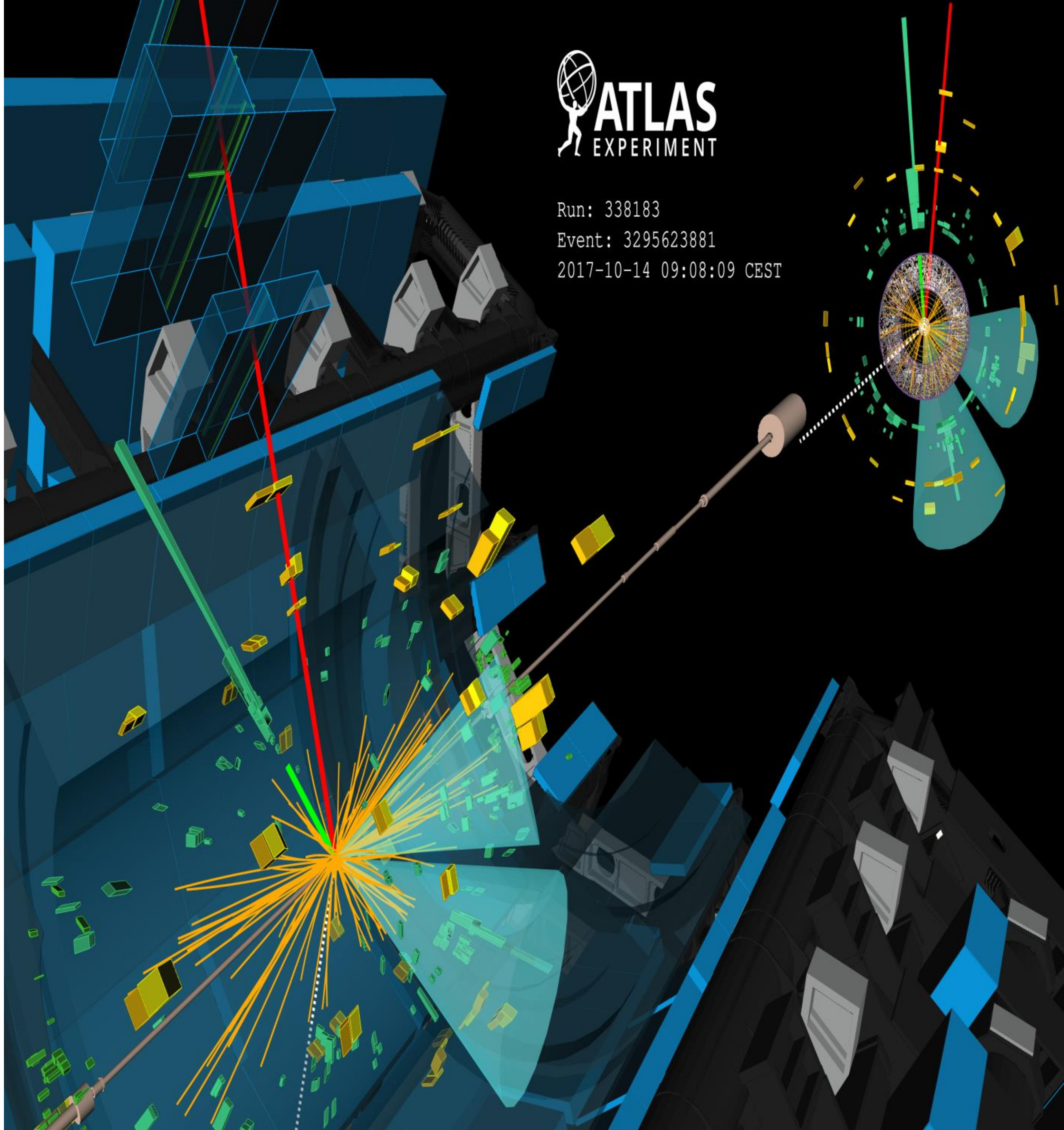


Table of contents

Top quark

What is it?

m = 172.52 GeV/c²

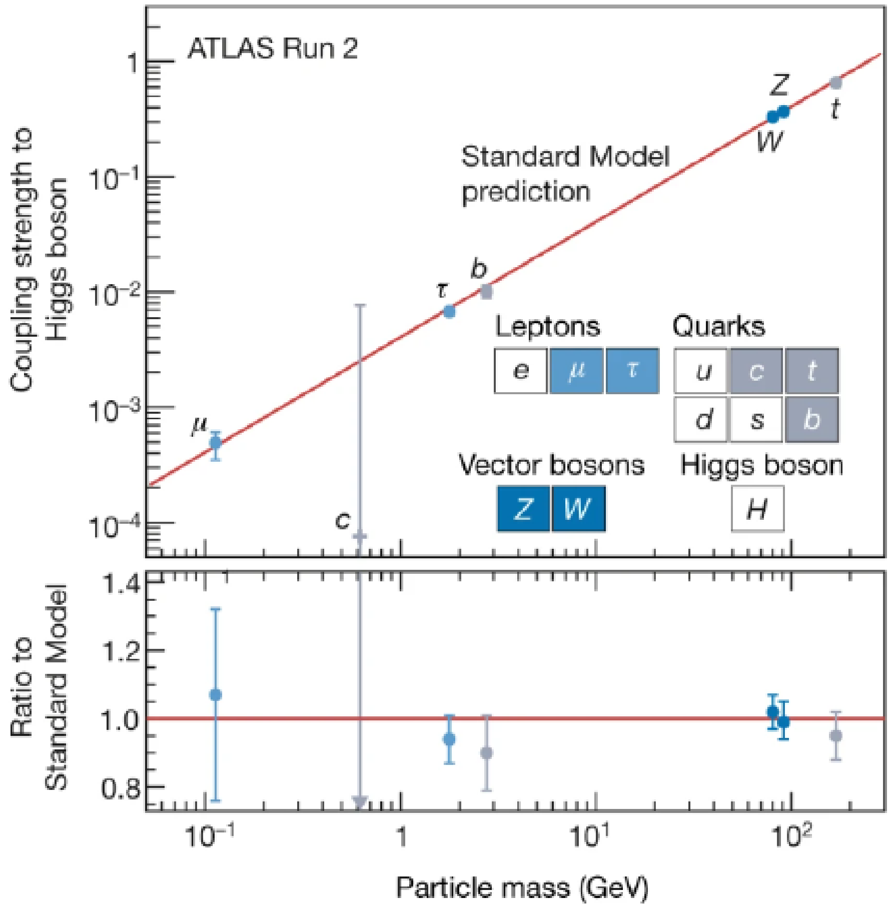
q = +2/3

t

top

The most massive quark, 40 times more than the previous one (b quark); strongly coupled to Higgs boson

[Nature 6075 52-59 (2022)]



$$\tau_t \sim 10^{-25} s < \text{ hadronisation time}$$

Phys. Rev. D 110, 030001 (2024)

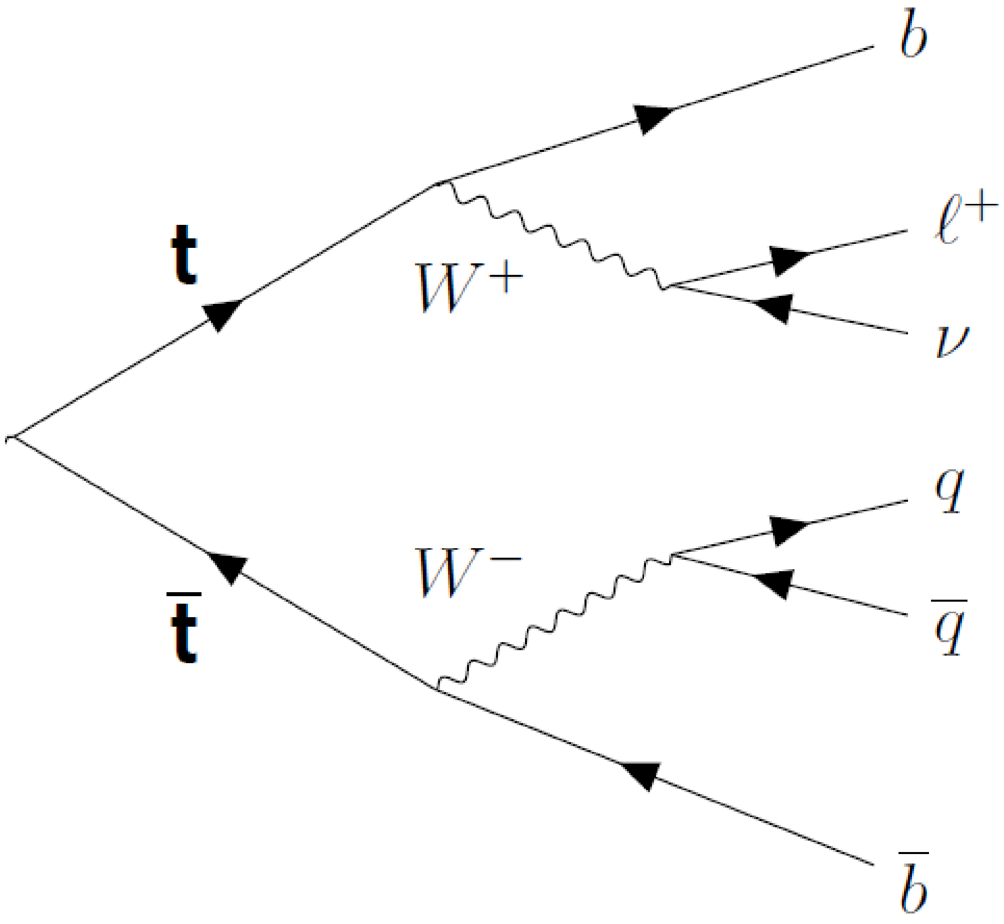
Top quark

How does the top decay?

$$\tau_t \sim 10^{-25} s < \text{ hadronisation time}$$

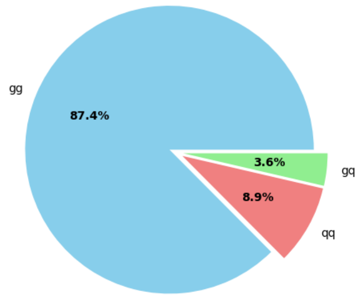
Phys. Rev. D 110, 030001 (2024)

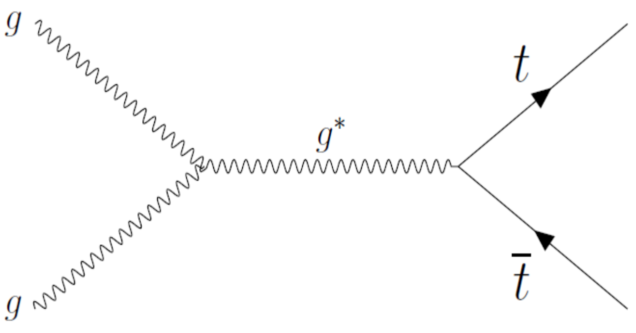
- Semi-Leptonic (1L) decay: one W decays hadronically, the other leptonically
- Detector signature: ≥ 4 jets + MET + 1 lepton
- Leptonic t and Hadronic t

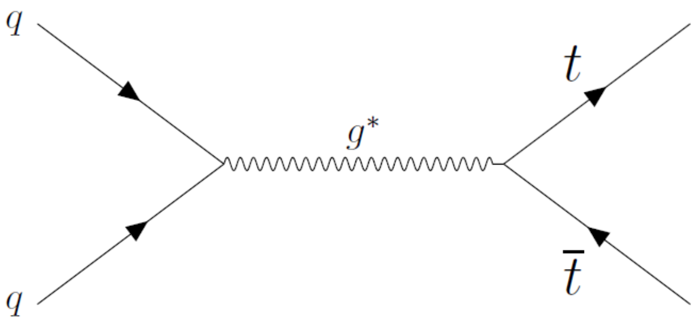


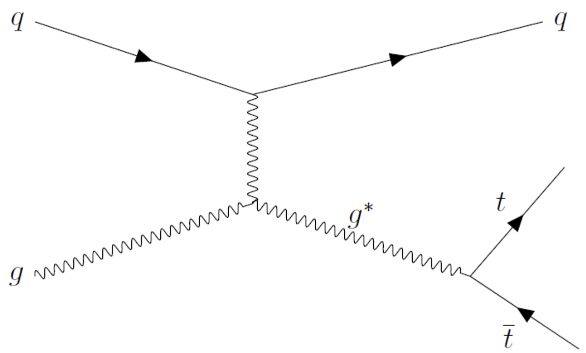
Top quark

How is top produced?









gg channel

Most abundant one

10% attractive colour singlet $^{1}S_0^{[1]}$

90% repulsive colour octet $^{1}S_0^{[8]}$

qq channel

Final state similar to gg

100% repulsive colour octet $^{3}S_0^{[8]}$

gq channel

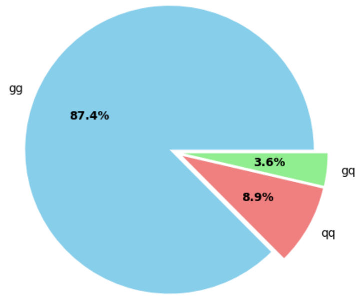
Different final state, more jets

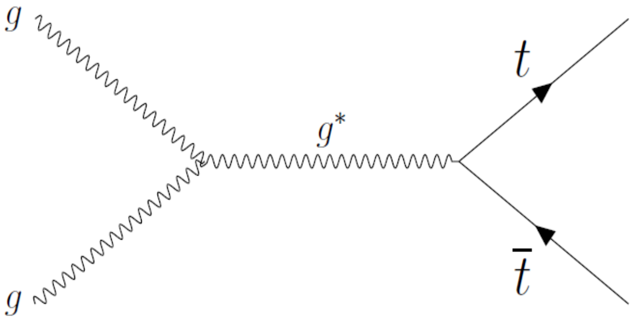
Partially colour singlet → mixed

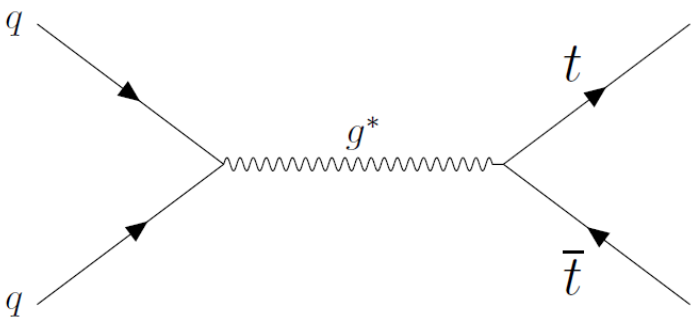
K. Behr- Hamburg International Summer School 2025

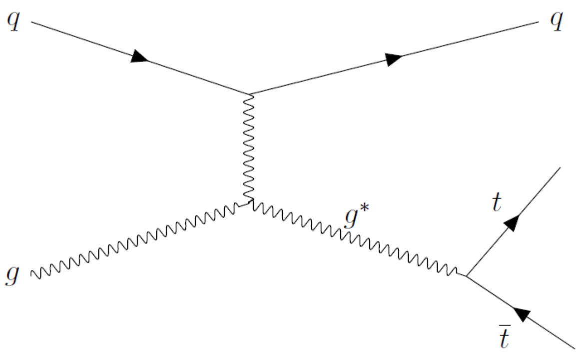
Top quark

How is top produced?









gg channel

Most abundant one

10% attractive colour singlet $^1S_0^{[1]}$

90% repulsive colour octet $^1S_0^{[8]}$

qq channel

Final state similar to gg

100% repulsive colour octet $^3S_0^{[8]}$

gq channel

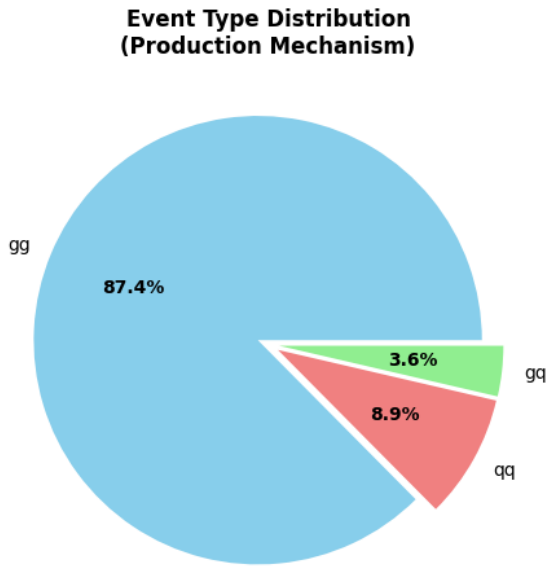
Different final state, more jets

Partially colour singlet → mixed

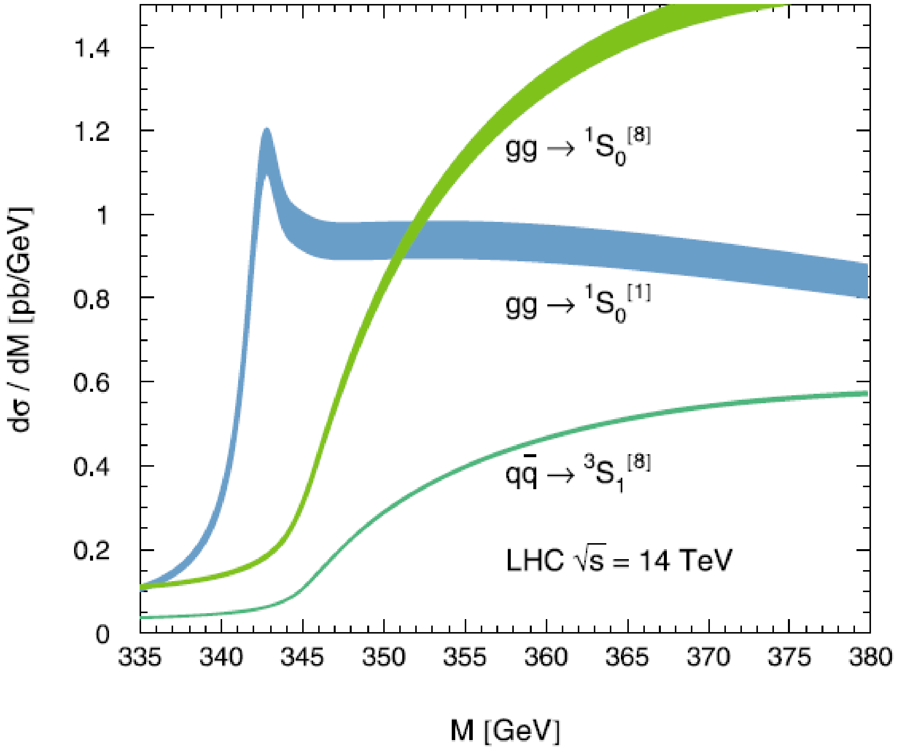
K. Behr- Hamburg International Summer School 2025

Top quark

Attractive colour singlet



Kiyo et al: EPJC 60(2009) 375-386



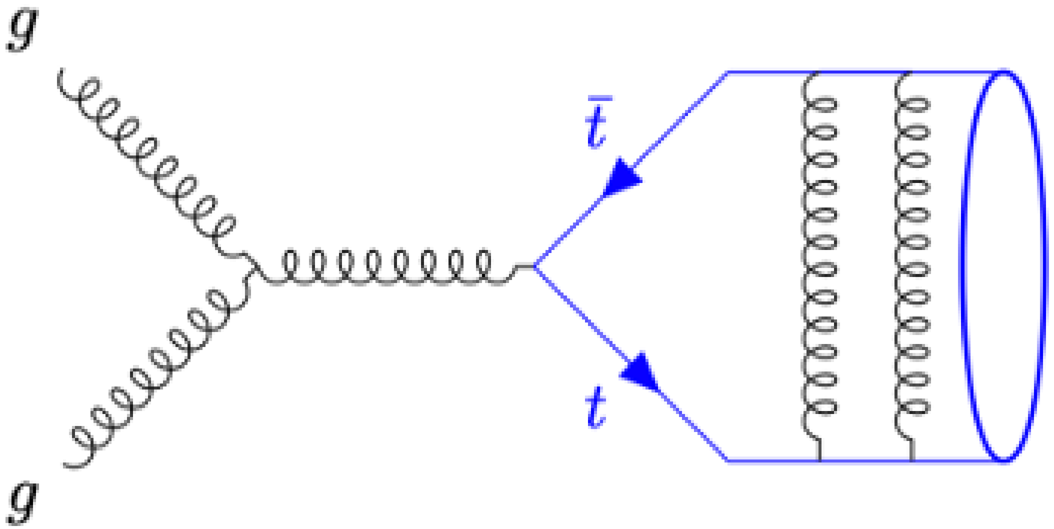
- Attractive colour singlet ${}^1S_0^{[1]}$ is the good candidate for a quasi-bounded state in the threshold region
- First predicted in 1987 (before top observation)
- Threshold region: $\mathrm{M} = 2 \cdot \mathrm{m_t}$

GOAL OF THIS PROJECT:
Classify the different colour states using properties of the production mode

Top quark

Quasi-bounded state

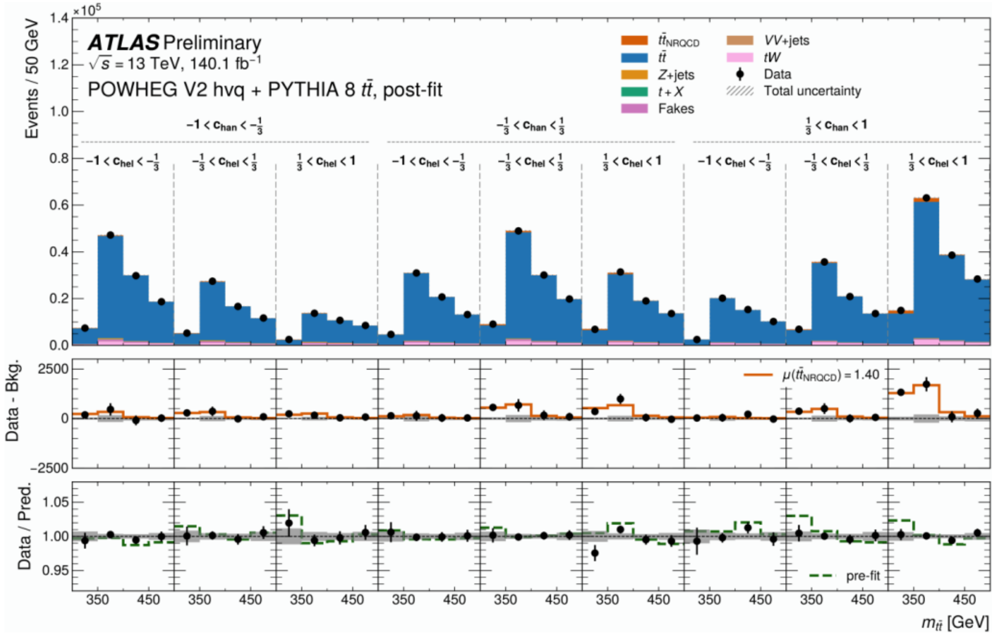
- Multiple gluon exchanges between t and tbar
- Creation of a quasi-bounded state, called “toponium”
- Mass just below the production threshold



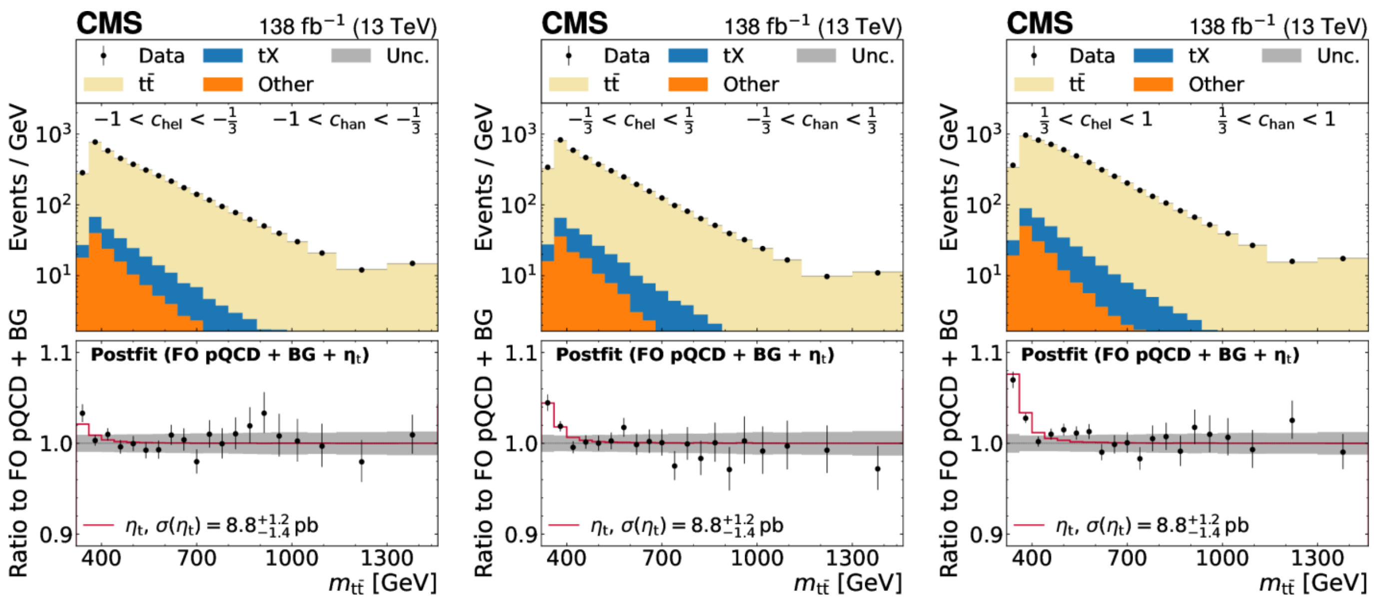
Top quark

Quasi-bounded state - observation

Never thought of observing it at the LHC due to low cross-section and detector resolution.
But, in 2025, it was observed, with an inclusive measurement:



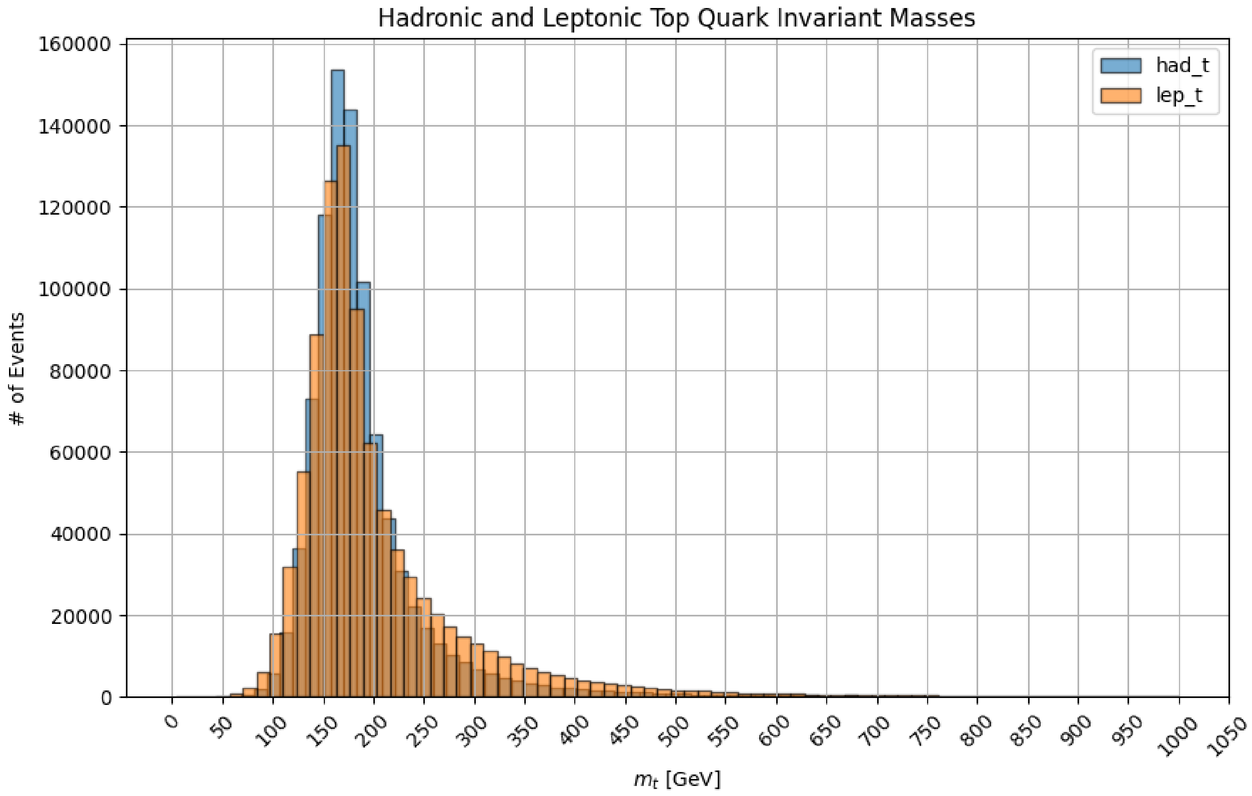
ATLAS-CONF-2025-008



CMS-TOP-24-007

Dataset

What is our starting point?



- MC simulated ttbar data, after ttbar reconstruction from SPANet
- The production channel of each event is known
- We compute four vectors of all final state particles

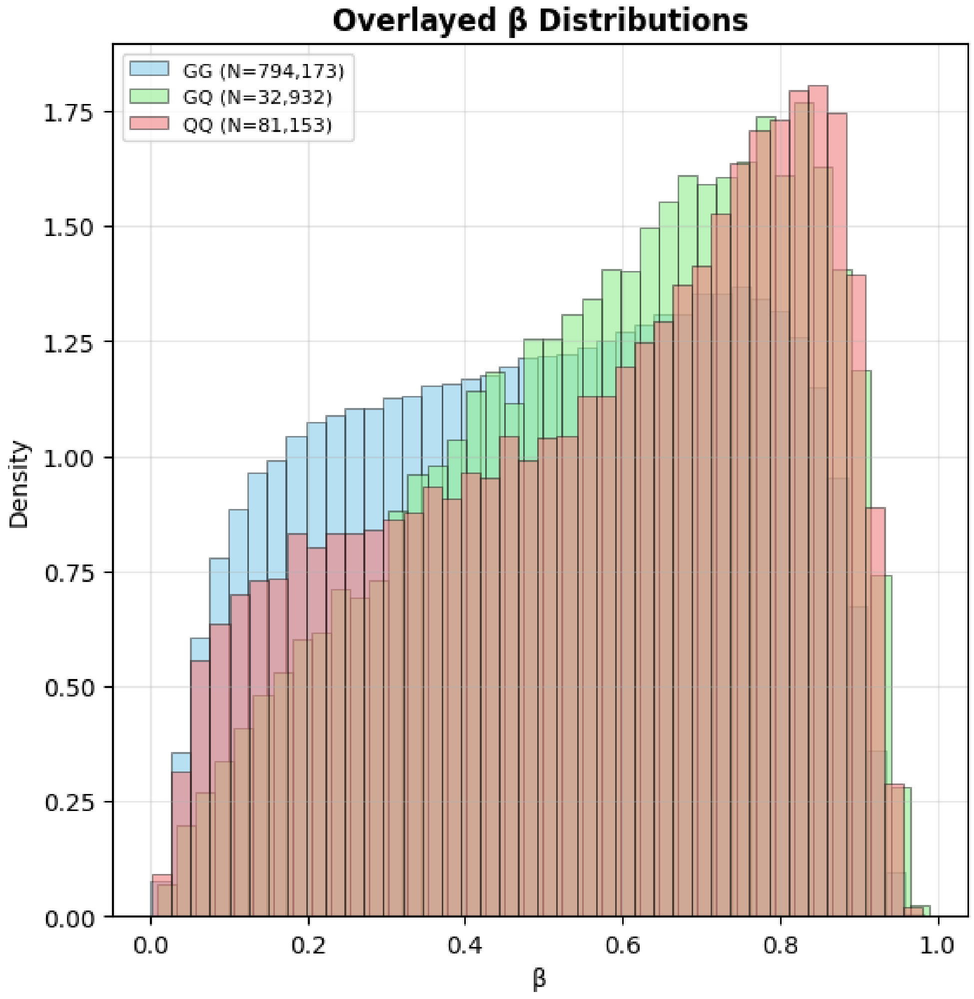
Dataset

Building needed bricks

- Computing high-level variables:
  - p_T for $t\bar{t}$
  - $\Delta\eta$, $\Delta\mathrm{R}$ for $t\bar{t}$ and $\mathrm{b}\bar{\mathrm{b}}$
  - β
  - $|\beta_{\mathrm{z}}|$

- Variables usually involved in cut-based analysis

- Good discriminating power



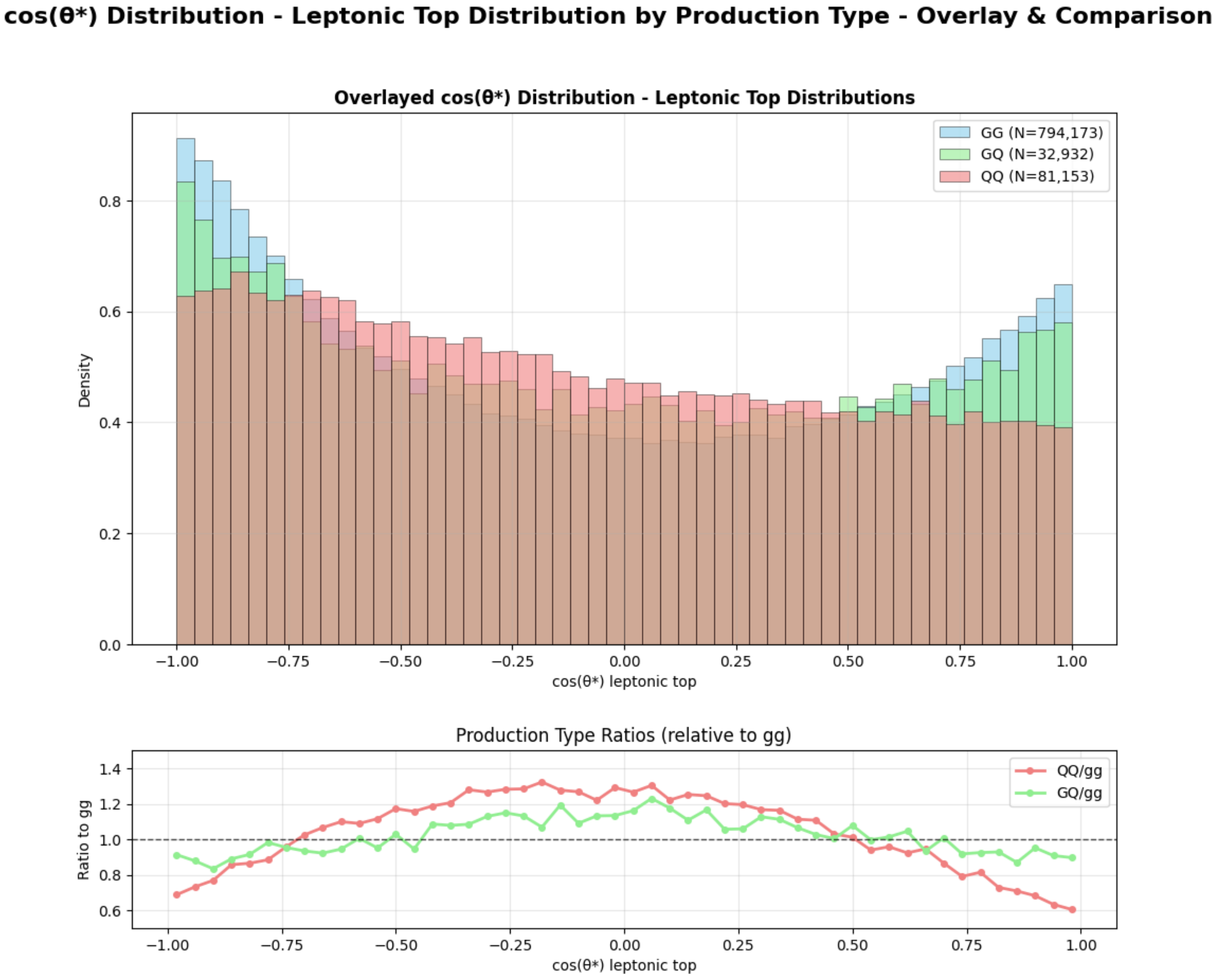
Dataset

Angular variables

Also computed angular variables:

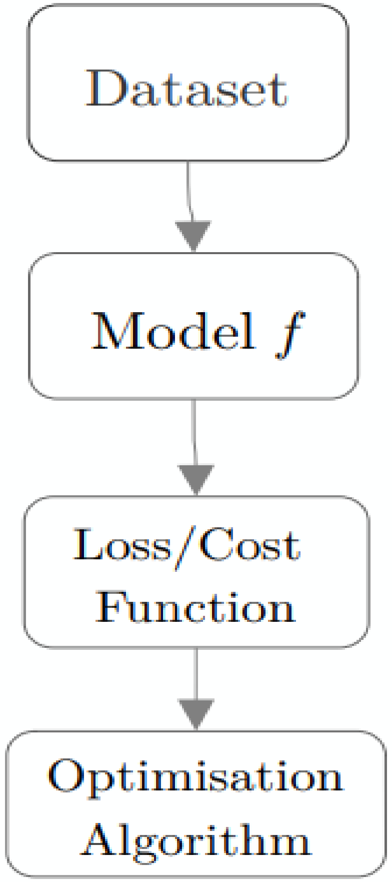
- $\cos\varphi_{\text{ij}} = \text{chel}$, angle between lepton and d-quark
- chan, same but with z flip of lepton
- $\cos\theta*$, angle between lep/had t and flight direction of ttbar system

Through them, information about spin and angular distribution



Deep Neural Network

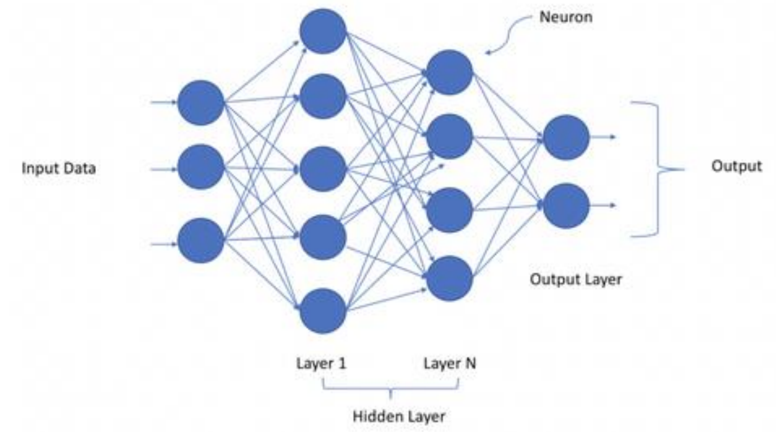
What is it?



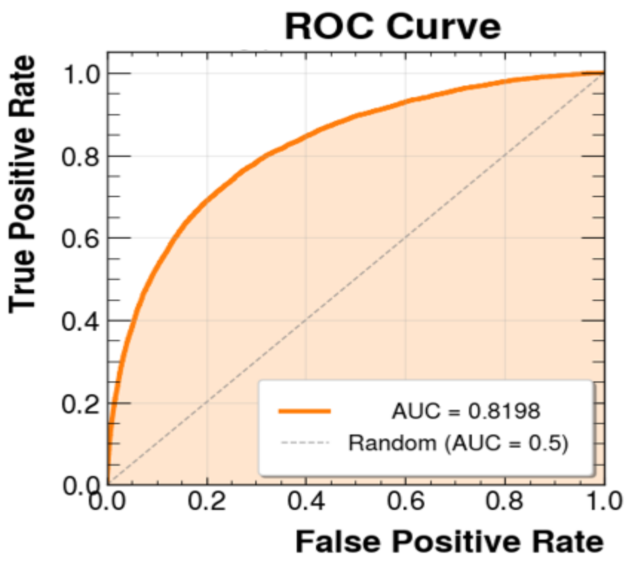
S. Jiggins - DESY Summer Lecture 2025

In this study:

- **LOSS**: Binary and categorical crossentropy
- **METRIC**: weighted accuracy
- **13 FEATURES**: β, $|\beta_z|$, $m_{t\bar{t}}$, $p_{T\,t\bar{t}}$, *jets_per_event*, $\Delta\eta_{\,b\bar{b}}$, $\Delta R_{\,b\bar{b}}$, $\Delta\eta_{\,t\bar{t}}$, $\Delta R_{\,t\bar{t}}$, $\cos\varphi$, c_han , $\cos\theta\,{*}_{had}$, $\cos\theta\,{*}_{lep}$
- **DATASET SPLIT**: (training, validation, testing) = (60:20:20)

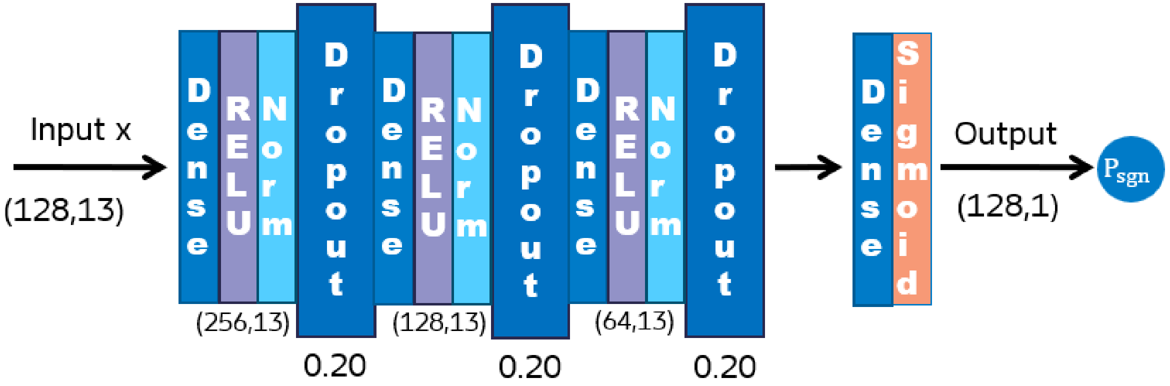


T. Lesort. Continual Learning: Tackling Catastrophic Forgetting in DNNs

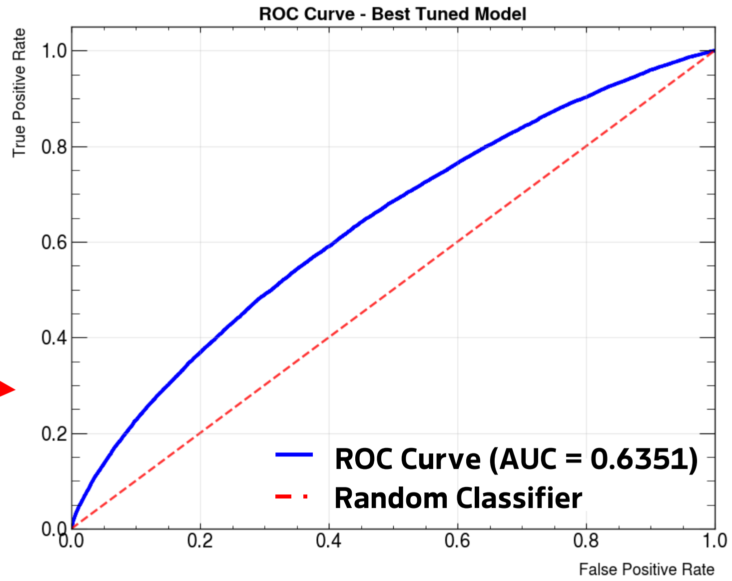


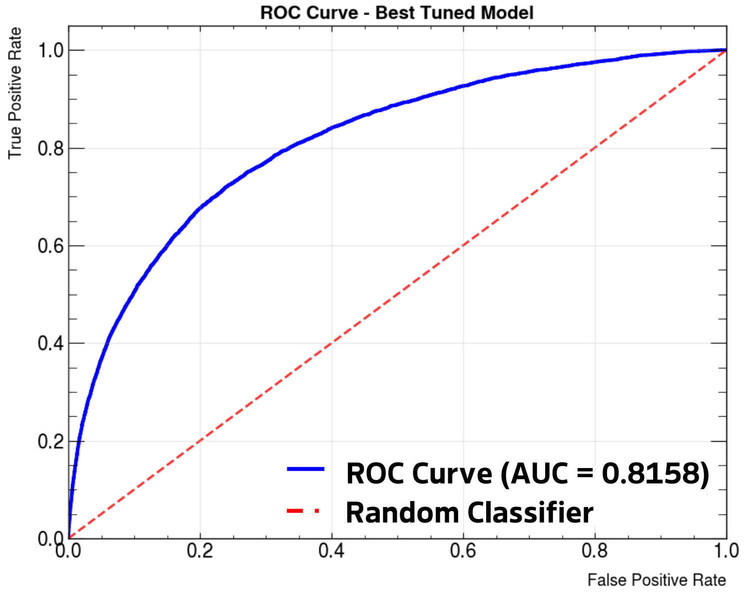
DNN implementation

Binary classification – separate analysis



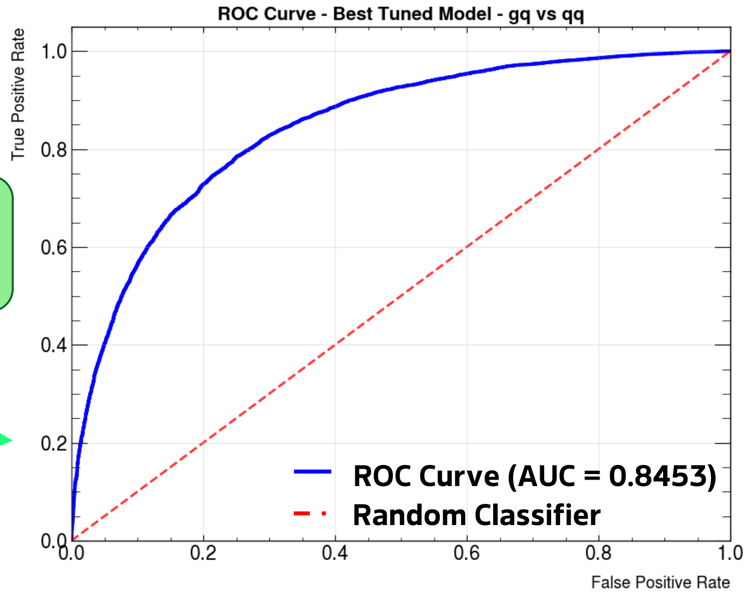
gg vs qq classifier
Not as good as
other channels!





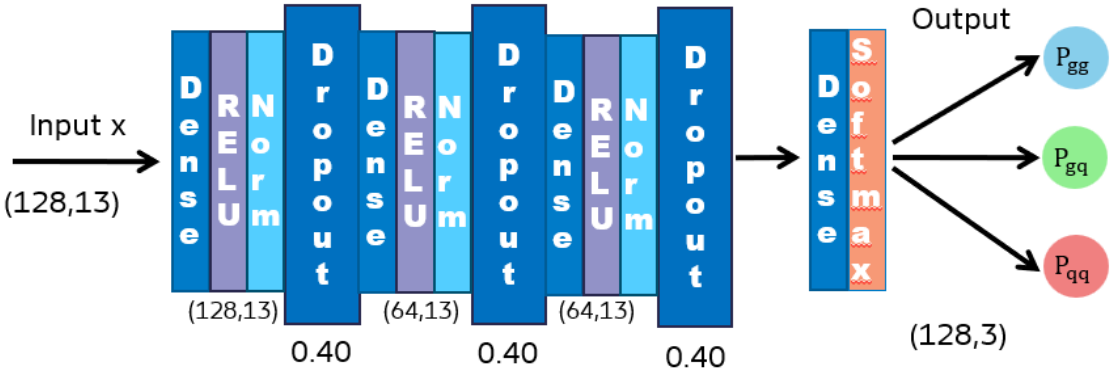
gg vs gq classifier
Good performance

gq vs qq classifier
Good performance



DNN implementation

3 classes classification – [gg,gq,qq]

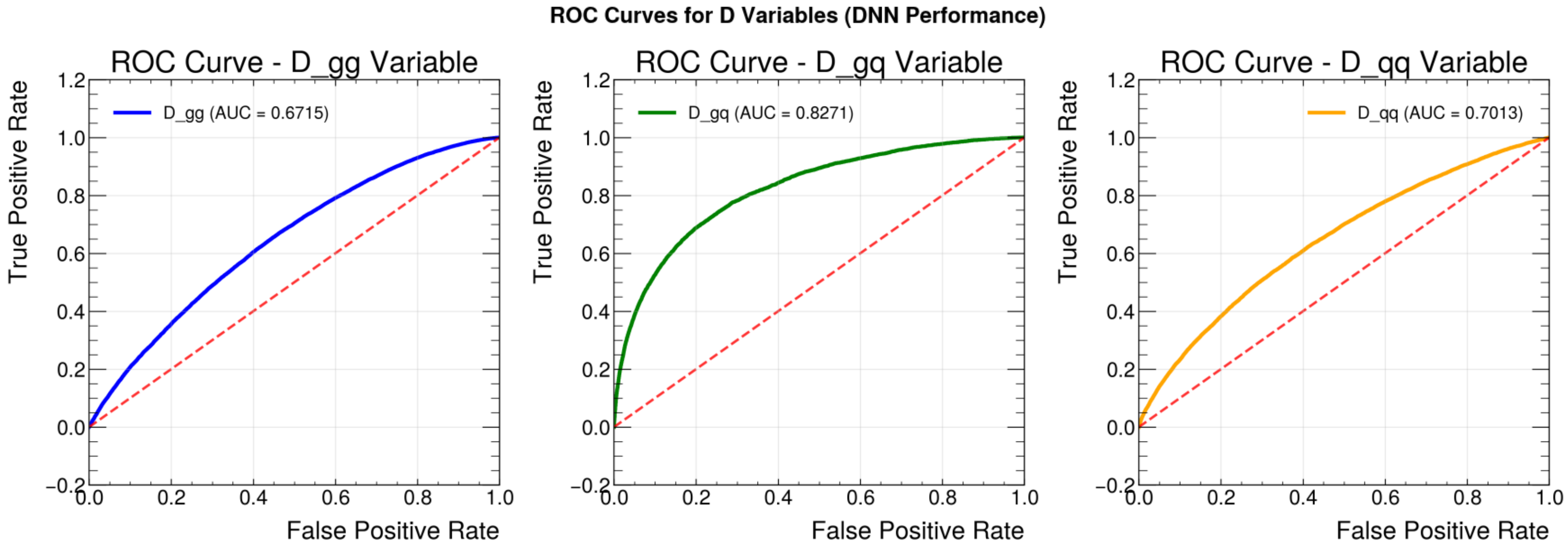


Discriminator is used to evaluate model's performance:

$$D_{gg} = \frac{P_{gg}}{f_{gq}P_{gq} + f_{qq}P_{qq}}$$

$$D_{gq} = \frac{P_{gq}}{f_{gg}P_{gg} + f_{qq}P_{qq}}$$

$$D_{qq} = \frac{P_{qq}}{f_{gg}P_{gg} + f_{gq}P_{gq}}$$



DNN implementation

Feature ranking and optimisations

- **FEATURE ANALYSIS:**

Gradient-Based Feature Importance by Tensorflow,
→ ttbar_pT most discriminant variable

- **OPTIMISATIONS**:

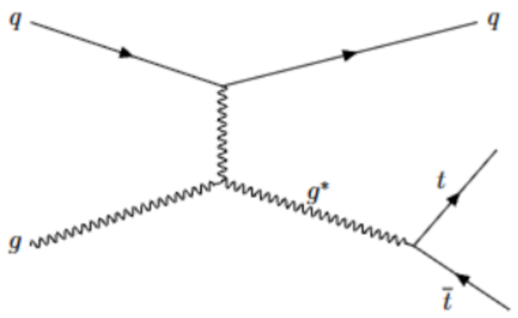
Testing different hyperparameters and choosing the configuration with the maximum AUC

ROC Curve computed over single variables distributions: DNN is always performing better

Variable	gg vs gq	gg vs qq	gq vs qq
Beta	59.7 %	57.4 %	58.1 %
Beta_z	51.3 %	57.9 %	56.1 %
ttbar pT	77.0 %	56.5 %	53.2 %
DNN 1 node	81.6 %	63.5 %	84.5 %

Conclusions

Could this help with toponium?



- gq channel is the one we can identify better, both in the 1 node and in the 3 nodes analysis
- DNN could be more performant than classical cut-based analysis
- It is worth investigating the production modes of ttbar to be able to probe the threshold region more deeply: <u>suppressing qq and gq production provides a cleaner data sample</u> for studies of *toponium* and also other measurements, such as *quantum entanglement*

- <u>*Future outlook:*</u>
  - *Increasing statistics*: from 900K to 10M events
  - Implementing *more complex architectures* (e.g. Transformers), which would allow us to extract potential additional features and difference from low-level variables, such as the jet 4-vectors

Thank you

<u>Thanks to:</u>

K. Behr, E. Jones, M. Gonzalez , F. Jolly for the guidance

ATLAS group for the hospitality

Contacts
Valerio Tinari

E-Mail:
valerio.tinari@desy.de
tinari.1998628@studenti.uniroma1.it

Deutsches Elektronen-Synchrotron DESY
ATLAS Group
Bldg. 01c | O1.301
Notkestraße 85
22607 Hamburg
www.desy.de

BACKUP

Dataset

Kinematic variables computation

- Beta for ttbar system: $\beta_{t\bar{t}} = \frac{p_{t\bar{t}}}{E}$ → *boost of the system*
- Beta_z for ttbar system: $|\beta_{Z\,t\bar{t}}| = \frac{p_{Z\,t\bar{t}}}{E}$ → *boost of the system along z-axis*
- *Delta eta*: $\Delta\eta = \eta_t - \eta_{\bar{t}}$
- *Delta R*: $\Delta R = R_t - R_{\bar{t}}$

Dataset

Angular variables computation

- $\cos\varphi_{\mathrm{ij}}$ = chel : angle between lepton and down-type quark , after boosting them to top parent ref frame

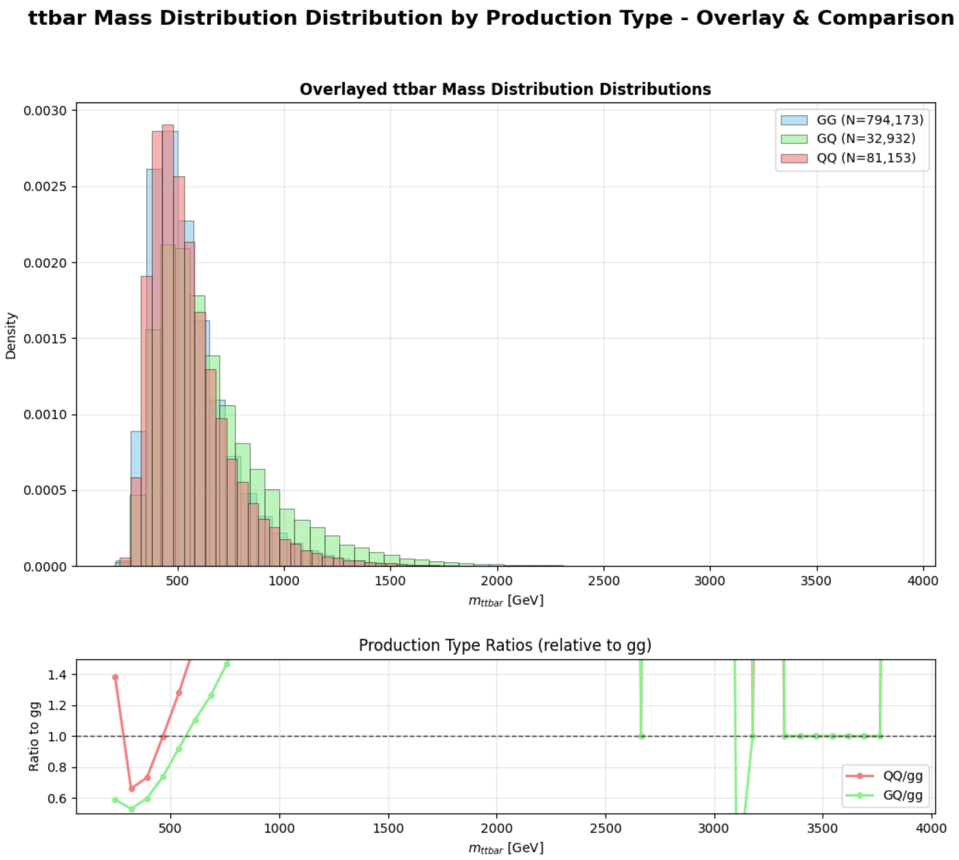
$$< \cos\varphi > = < \hat{\ell} \cdot \hat{d} >$$

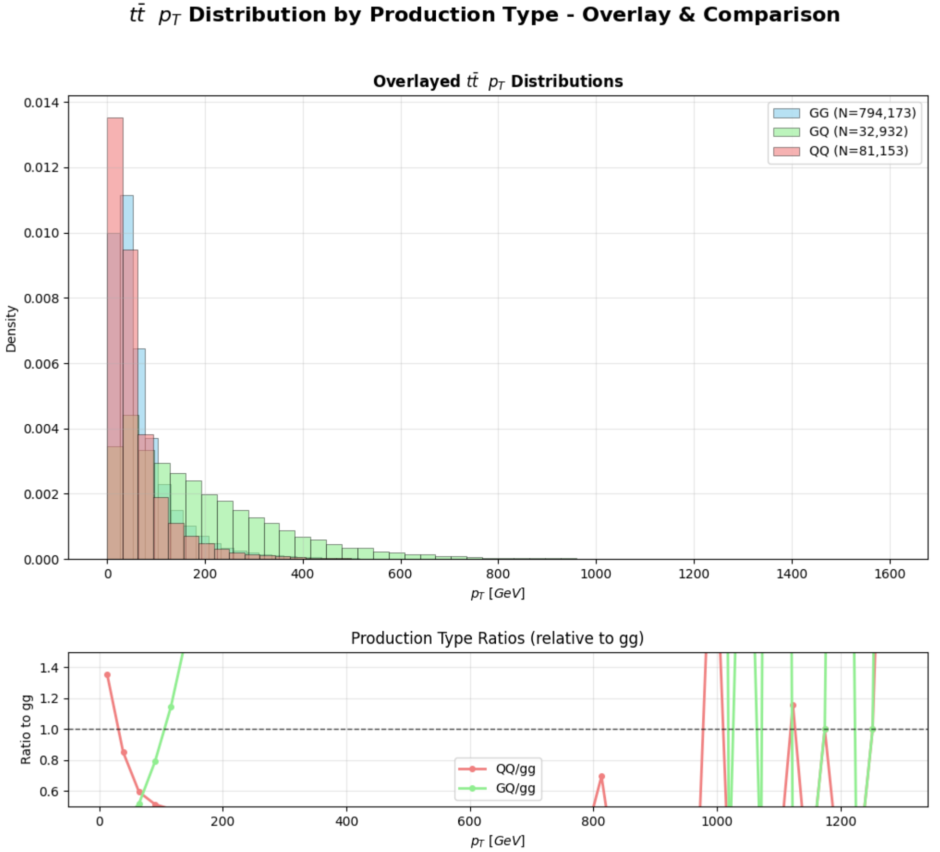
- Chan: the same of $\cos\varphi$, but with flip of the z-direction of lepton before the product

- $\cos\theta^{*}$: angle between one top and flight direction of ttbar

Dataset

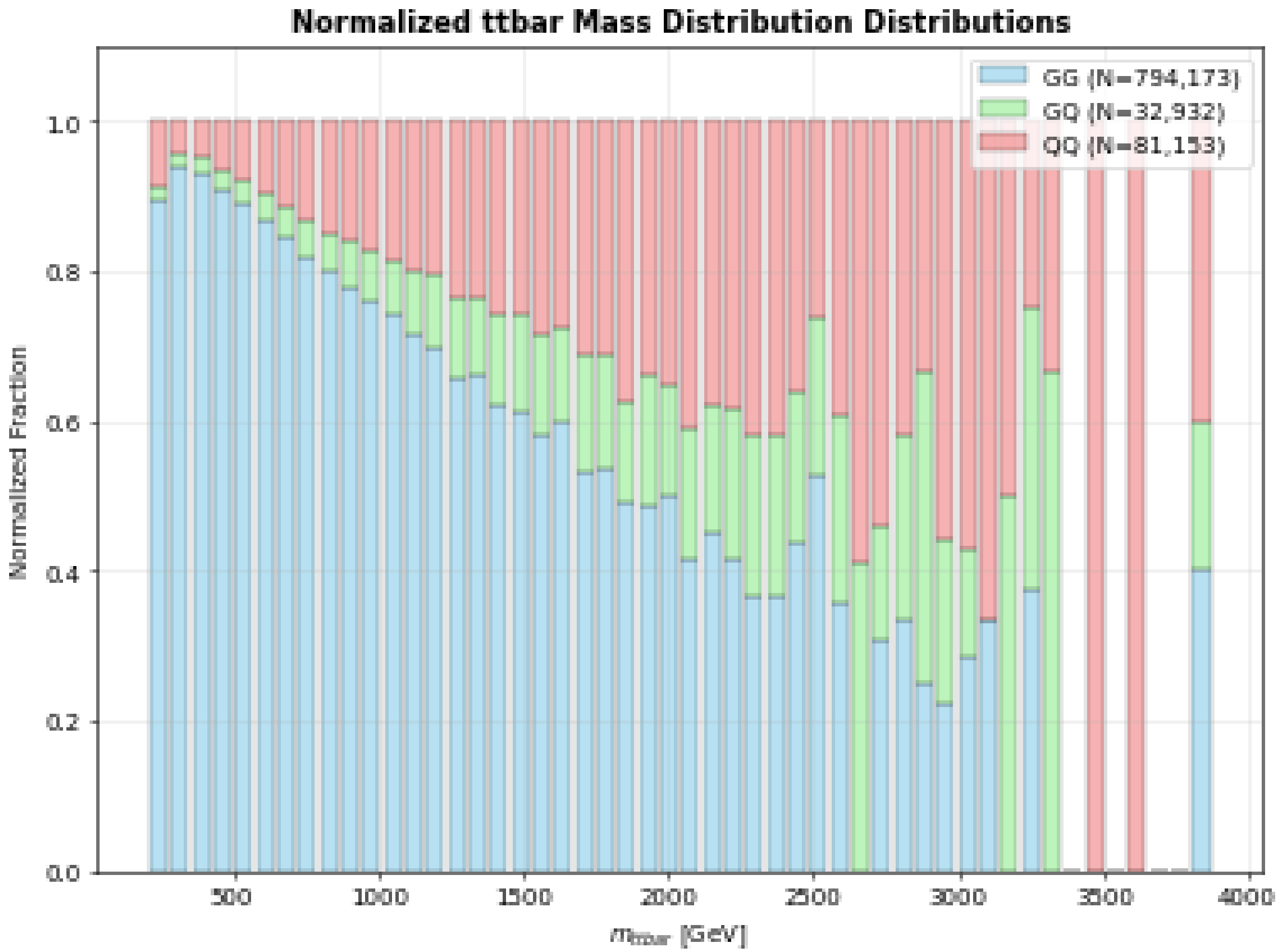
ttbar invariant mass and pT of ttbar





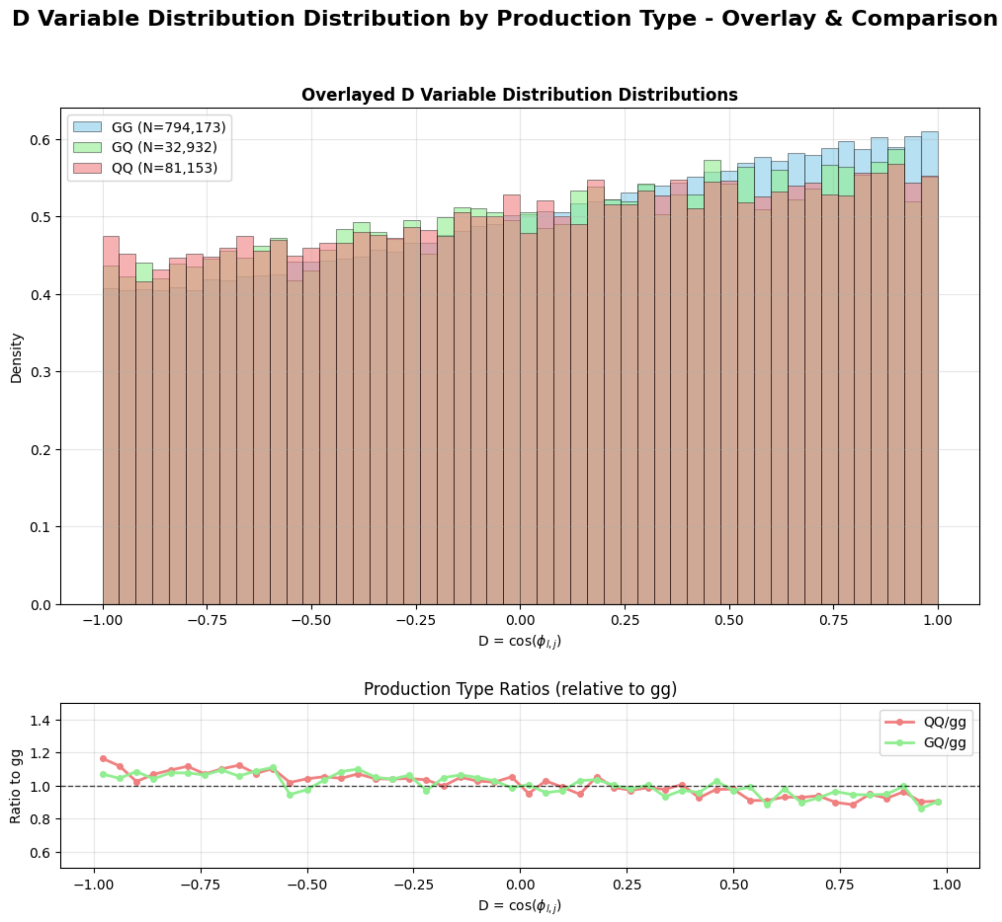
Dataset

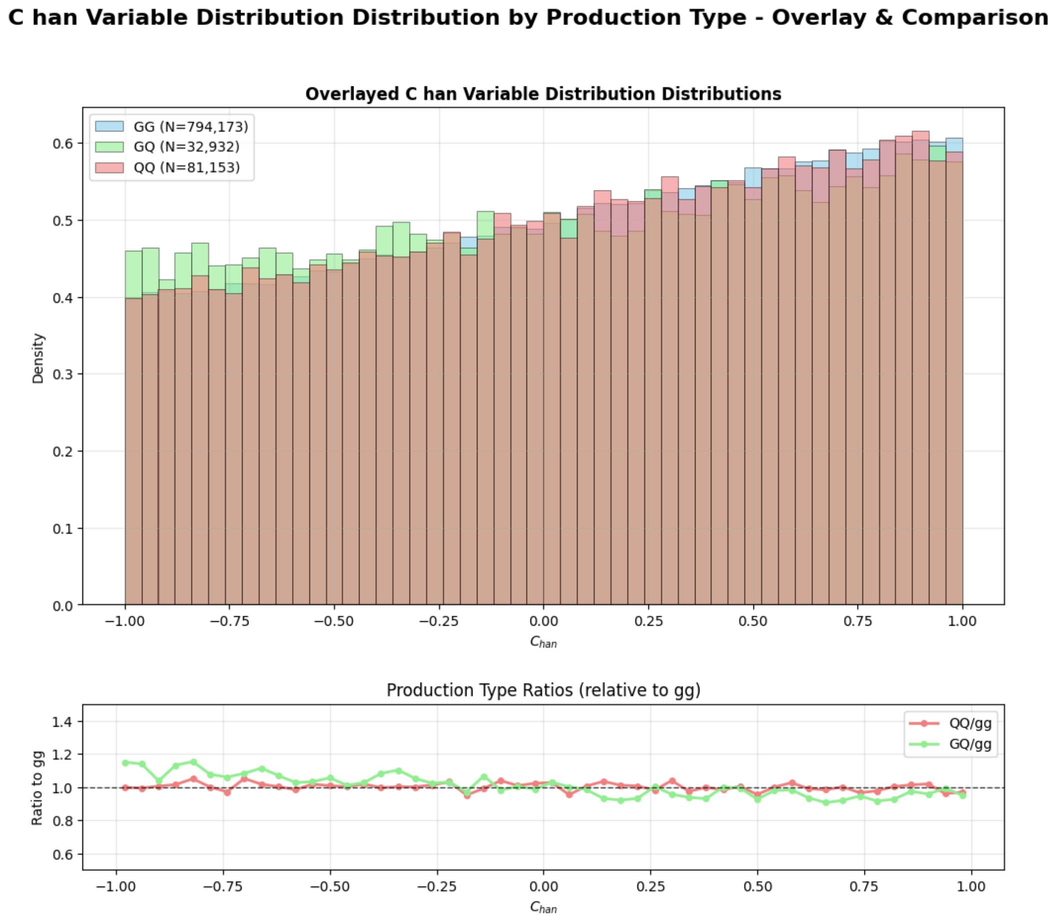
Normalized ttbar invariant mass per production mode



Dataset

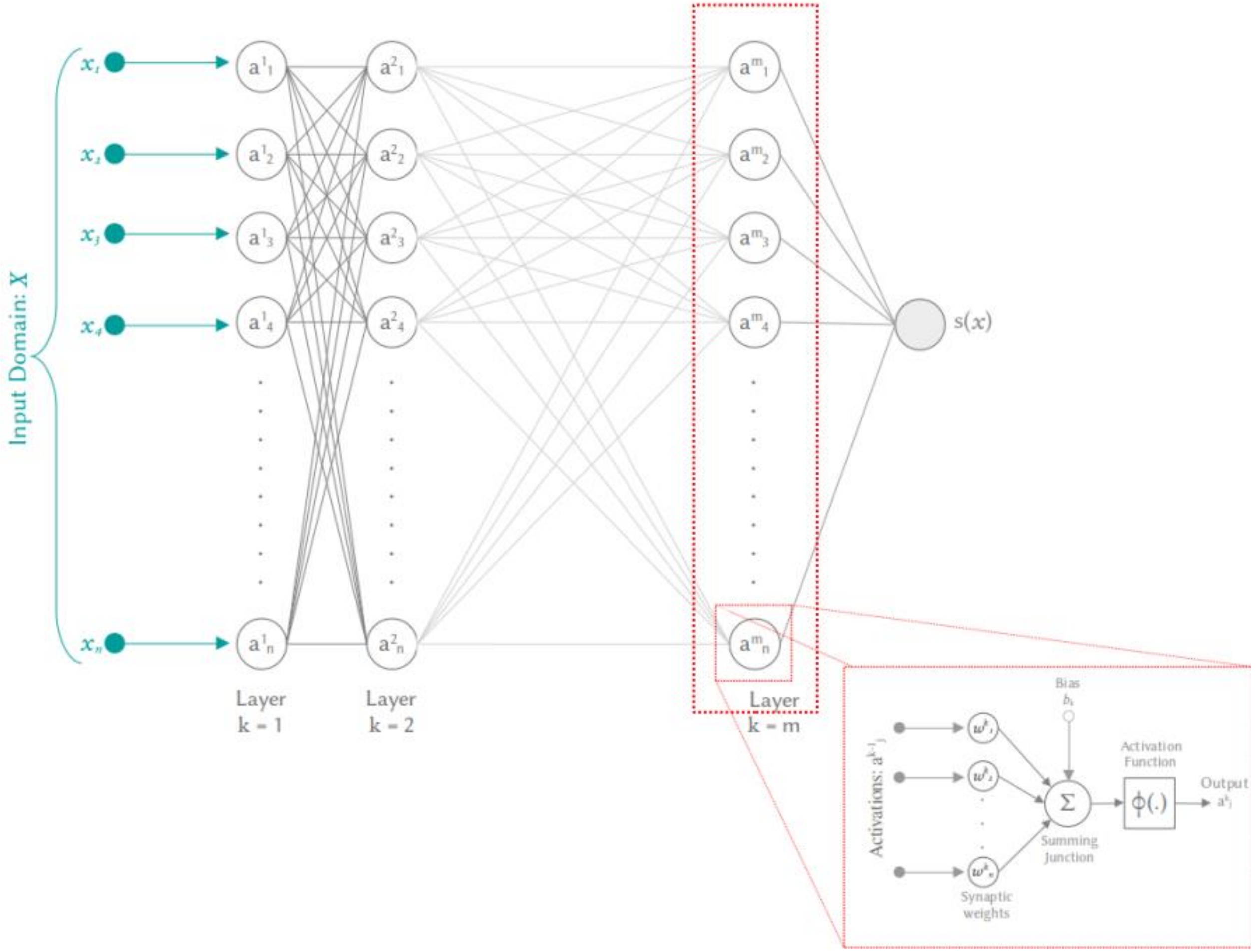
Angular variables per production mode





Deep Neural Network

What is it?



S. Jiggins - DESY Summer Lecture 2025

- Neuron structure:

$$a_i^{(k+1)} = \phi\left(\sum_j^n w_{i,j}^{(k)} a_j^{(k)} + b^{(k)}\right)$$

- Neural Network:

$$\begin{bmatrix} a_0^{(k+1)} \\ \dots \\ a_n^{(k+1)} \end{bmatrix} = \phi\left(\begin{bmatrix} w_0^{(0)} & \dots & w_n^{(0)} \\ \dots & \dots & \dots \\ w_0^{(k)} & & w_n^{(k)} \end{bmatrix}\begin{bmatrix} a_0^{(k)} \\ \dots \\ a_n^{(k)} \end{bmatrix} + \begin{bmatrix} b_0^{(k)} \\ \dots \\ b_n^{(k)} \end{bmatrix}\right)$$

Binary and categorical crossentropy loss

- Binary crossentropy (BCE):

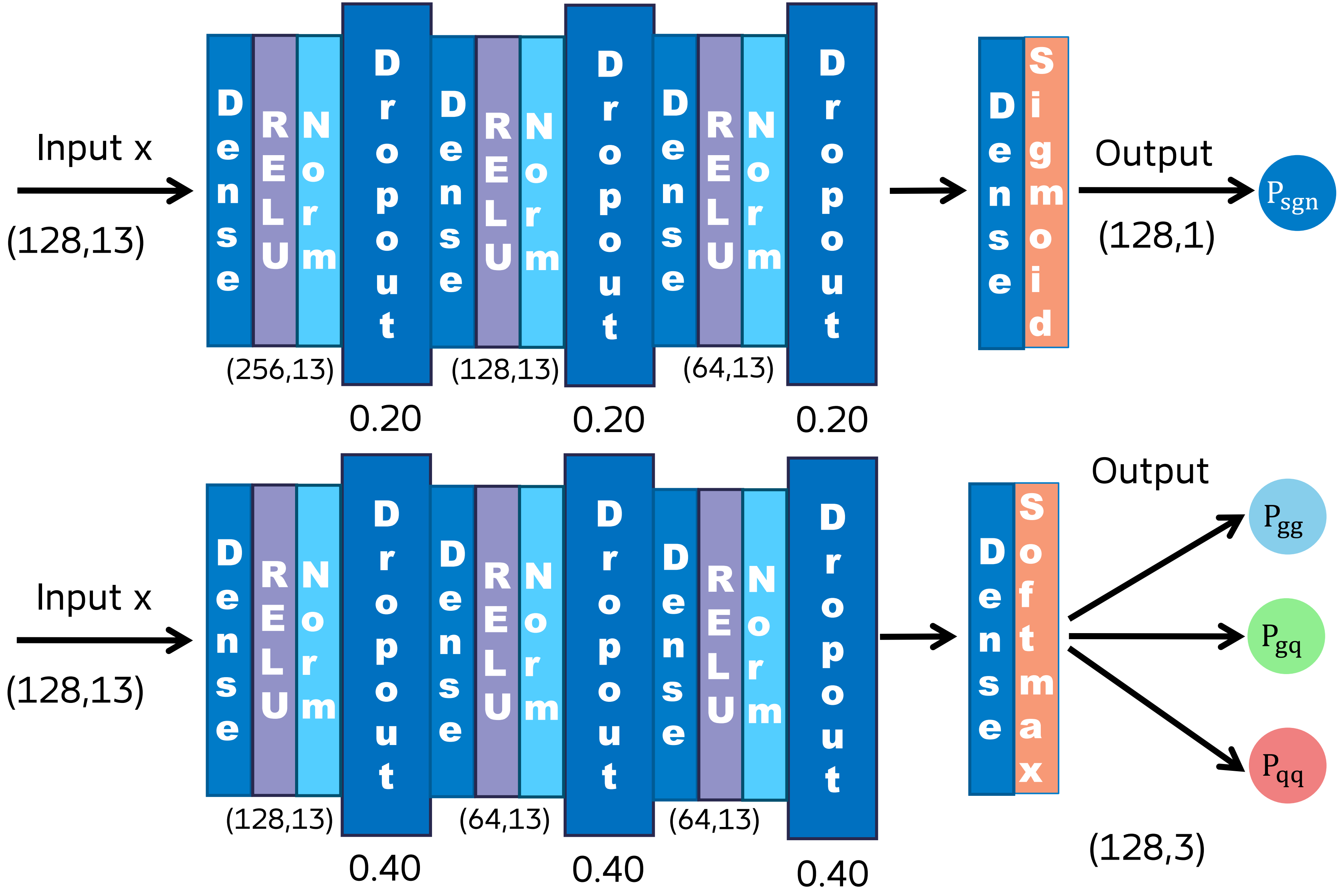
$$\mathrm{BCE} = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(p_i) + (1-y_i)\log(1-p_i)]$$

Binary Cross-Entropy measures the distance between the true labels and the predicted probabilities. When the predicted probability pipi is close to the actual label yiyi, the BCE value is low, indicating a good prediction.

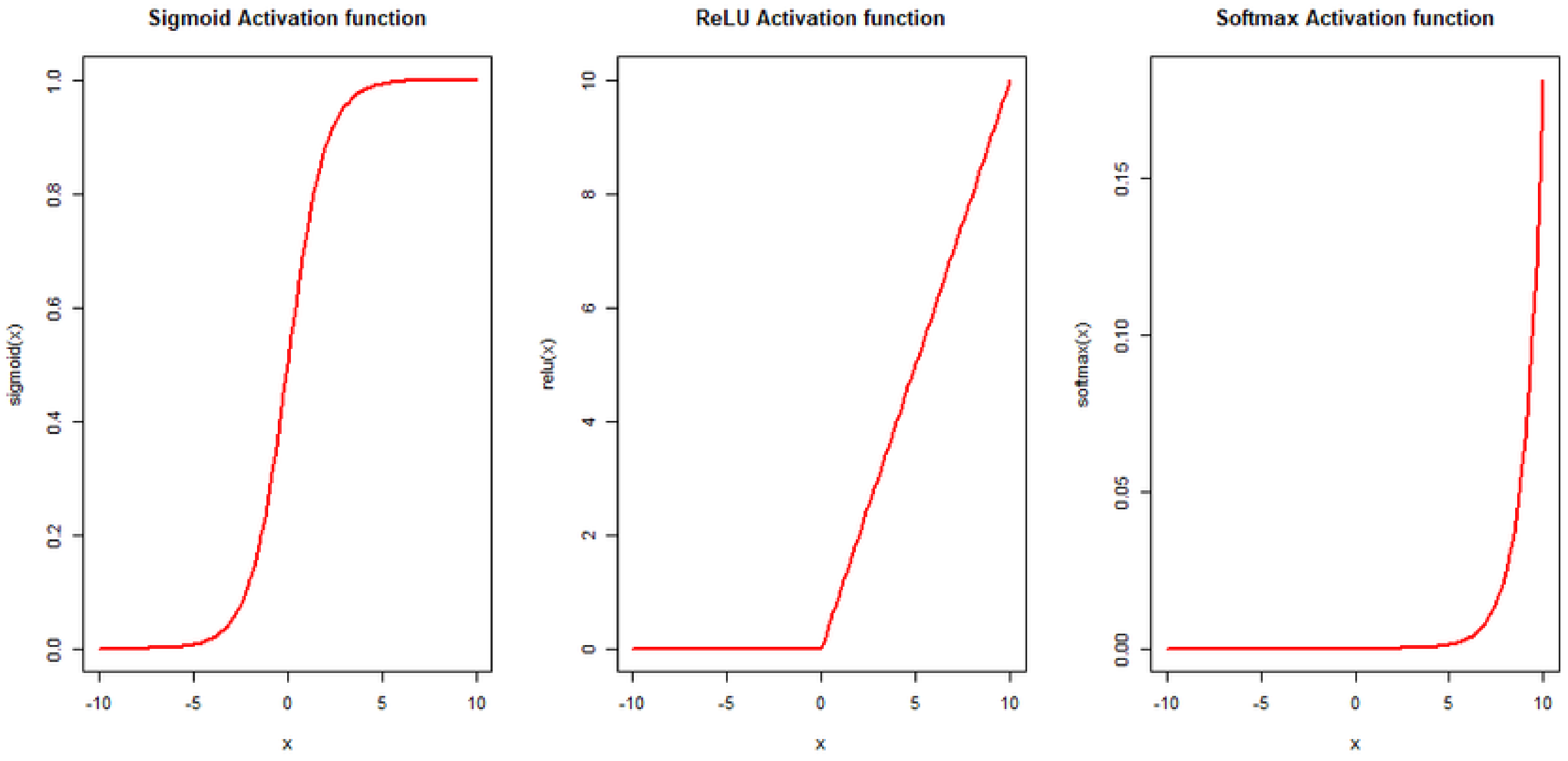
- Categorical crossentropy

$$L(y,\hat{y}) = -\sum_{i=1}^{C} y_i \log(\hat{y}_i)$$

Model architecture



Different activations



DNN implementation

Consistency checks

BDT confirms results from DNN

- BDT training to cross-check results
- ROC curve and AUC score over most important variables' distribution (Beta, Beta_z, ttbar pT)
- Run without angular variables → no major changes

Classical AUC	**gg vs gq**	**gg vs qq**	**gq vs qq**
BDT	80.9 %	83.8 %	63.7 %
Beta	59.7 %	57.4 %	58.1 %
Beta_z	51.3 %	57.9 %	56.1 %
ttbar pT	77.0 %	56.5 %	53.2 %

DNN is performing better

Discriminator AUC	**gg signal**	**gq signal**	**qq signal**
BDT 3 classes	64.3 %	81.9 %	64.5 %

BDT documentation

Features ranking – BDT (gg vs qq channel)