Search for Axion-Like Particles in FCNC Decays of the Top Quark within the ATLAS Experiment

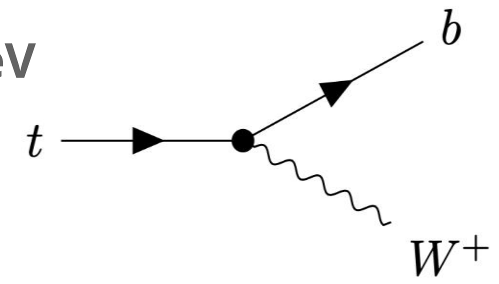
Martin Christiansen

Herbstschule HEP

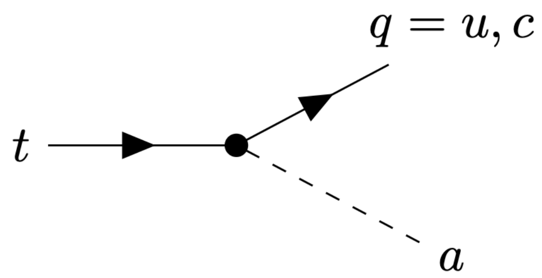
September 2026

Motivation

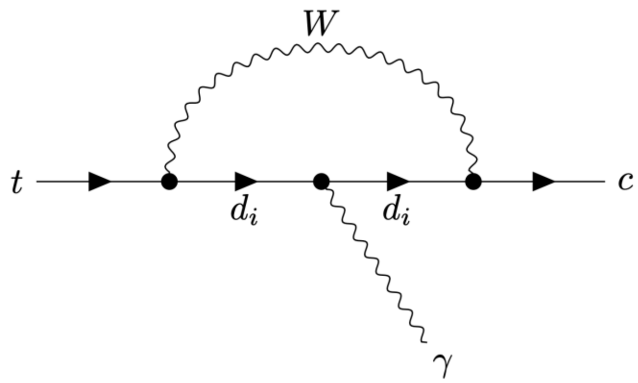
- Top quark is the heaviest particle in the Standard Model (SM), with $m_t \approx 173$ GeV
  - Decays almost exclusively via $t \rightarrow Wb$
  - Very short lifetime => decays before hadronisation



- Axion-like particles (ALPs) are pseudoscalar particles
  - appearing in many extensions of the SM
  - can induce flavour-changing neutral-current (FCNC) top-quark decays
    $t \rightarrow qa$, $q = u,c$



- In the SM, FCNC top-quark decays are highly suppressed
  => Observation would be a clear indication of new physics



$d_i = d, s, b$

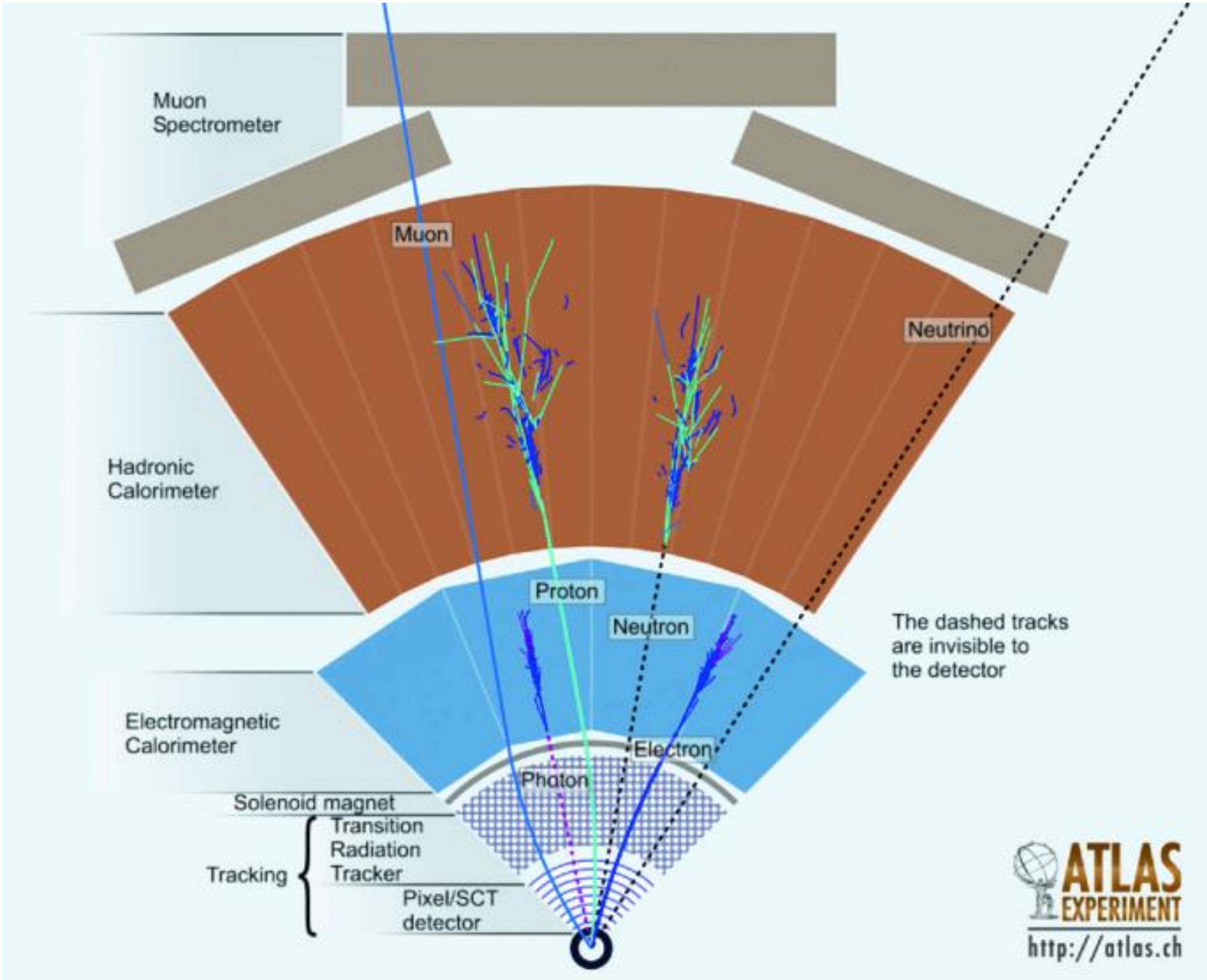
Experimental Signature

- Search for new physics via long-lived particles (LLPs)

- Hadronic ALP decay inside the calorimeter
  - ALP mass of 1-10 GeV => all decay products within a single jet
  - Little or no associated tracking activity
  - Large fraction of energy deposited in the hadronic calorimeter

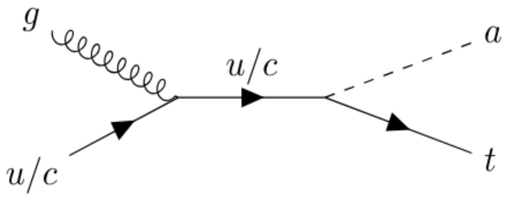
- Characteristic displaced-jet signature

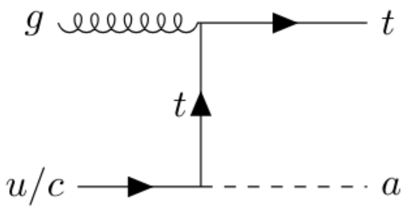
- Large $\log_{10}(E_{had} / E_{EM})$ value ("LogRatio" or "CalRatio")



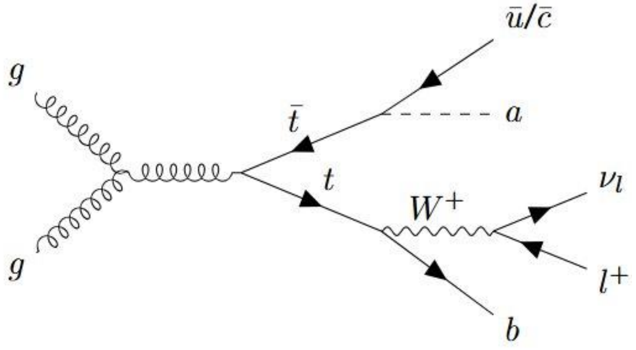
ALPs in Top-Quark Events

- Single Top + ALP
  FCNC production of a single top quark

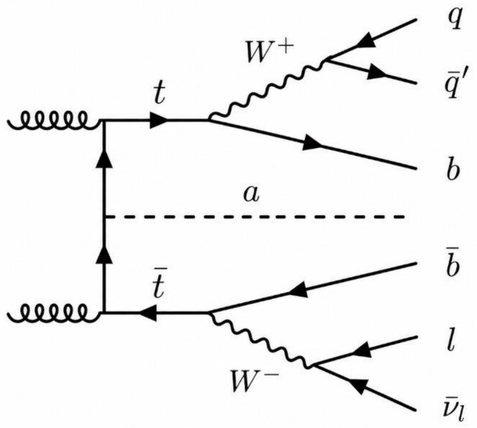




- Charming ALP
  FCNC top decay t → qa
  Focus of this talk



- $t\bar{t}$ + ALP
  Flavour-diagonal coupling



Event Reconstruction

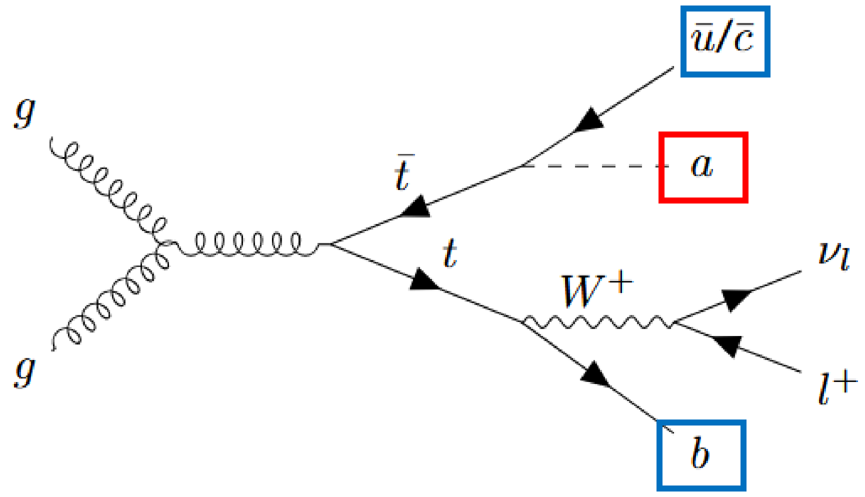
Signature: one W boson + at least three jets

Goal: Assign three jets

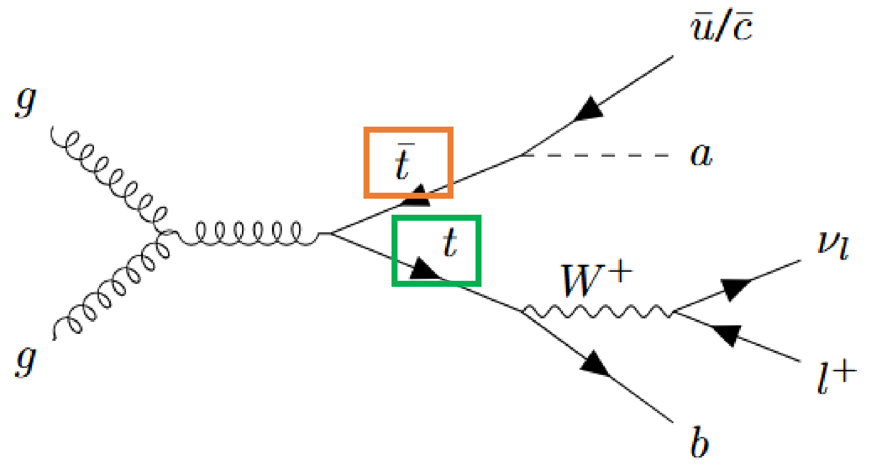
- **ALP candidate**: identified using LogRatio $\log_{10}(E_{had} / E_{EM})$
- **Remaining two jets**: likelihood approach
  - Joint probability density of observed jet kinematics under assignment hypothesis

$$L(\theta; m_1, m_2) = \prod_{i=1}^{2} f(m_i \mid \theta)$$

- Observables: four-vectors of the jets $\Rightarrow$ invariant masses of the reconstructed top quarks: $m_1 \equiv m_{t,aj}$ and $m_2 \equiv m_{t,Wj'}$
- Parameters θ: jet assignments of j, j' to u/c and b
- Calculate L for every combination $\Rightarrow$ choose maximum



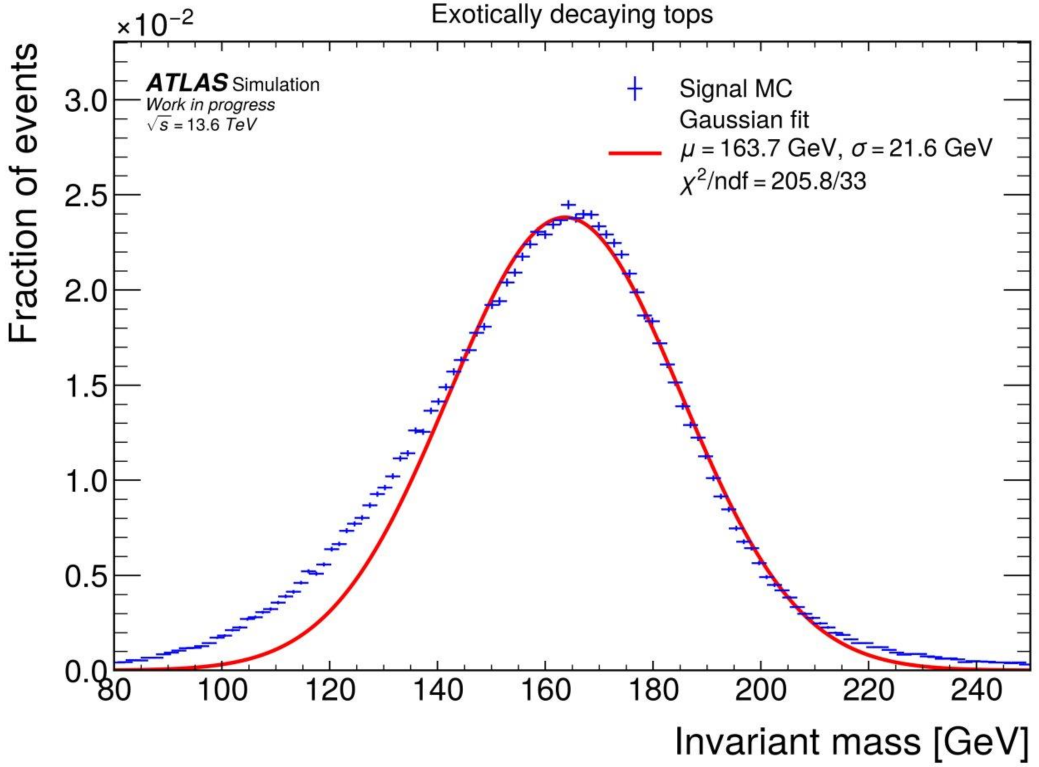
Event Reconstruction



Assume a Gaussian model for the reconstructed top mass

$$f(m|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(m-\mu)^2}{2\sigma^2}}$$

Parameters obtained from truth-matched simulation (separately for exotic and leptonic decays)



For each possible jet assignment:

- **reconstruct the top masses**
- **evaluate their compatibility with the expected mass distributions**
- **choose the assignment with the maximum log-likelihood (LLH = log L)**

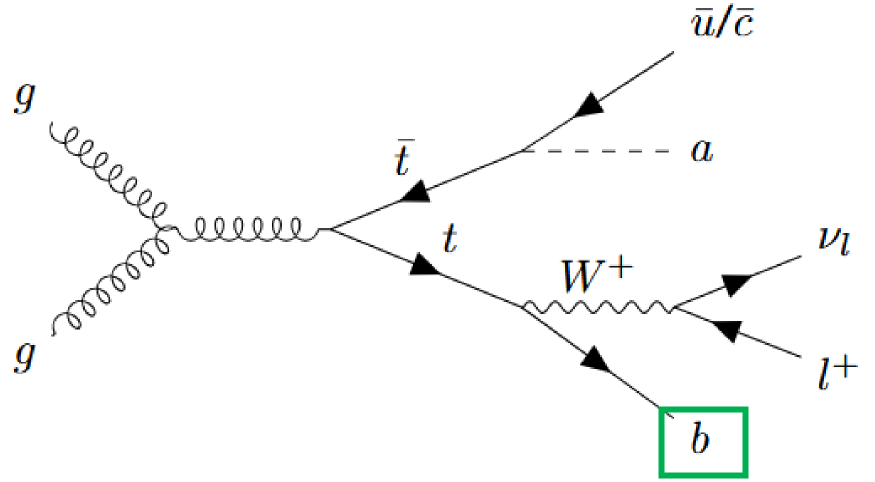
$$-\mathrm{LLH} = \frac{1}{2}\left(\frac{m_{t,aj} - \mu_{t,e}}{\sigma_e}\right)^2 + \frac{1}{2}\left(\frac{m_{t,Wj'} - \mu_{t,l}}{\sigma_l}\right)^2$$

Including b-tagging in the Likelihood

Choose two separate working points for the leptonic and exotic top

- Leptonic top: jet **expected to originate from b**
  Multiply L with $\epsilon_b \equiv P(\text{observed tag} \mid \text{b-jet})$
  e.g. for WP 85 %:
  - If observed tag is b: $\epsilon_b = \boxed{0.85}$
  - If observed tag is not b: $\epsilon_b = 0.15$

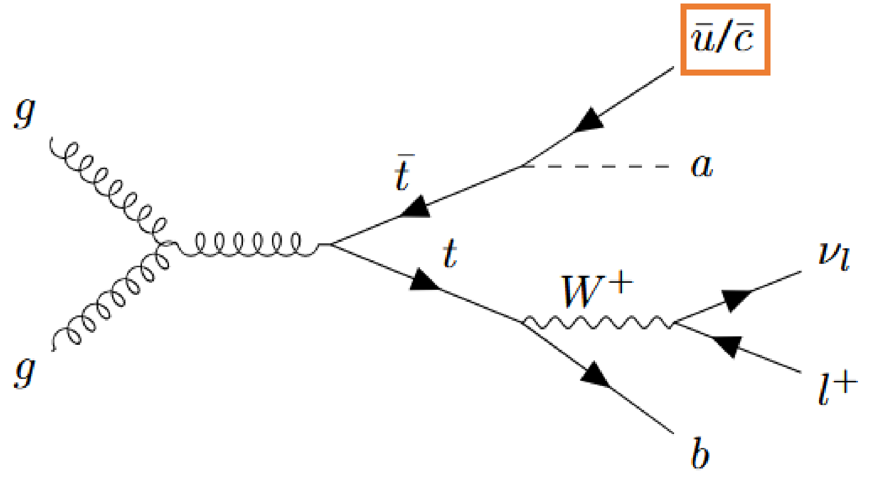
$$-\mathrm{LLH} = \frac{1}{2}\left(\frac{m_{t,aj} - \mu_{t,e}}{\sigma_e}\right)^2 + \frac{1}{2}\left(\frac{m_{t,Wj'} - \mu_{t,l}}{\sigma_l}\right)^2 - \log \epsilon_{u/c} \boxed{- \log \epsilon_b}$$



efficiency operating point	b-efficiency	c-efficiency	light-flavour efficiency
WP 65 %	0.66	0.00944	0.00026
WP 70 %	0.71	0.02081	0.00060
WP 77 %	0.78	0.05870	0.00218
WP 85 %	0.85	0.17505	0.01038
WP 90 %	0.90	0.33187	0.03237

The ATLAS Collaboration, Transforming jet flavour tagging at ATLAS. Nature Commun. 17 (2026) 541. http://dx.doi.org/10.1038/s41467-025-65059-6

Including b-tagging in the Likelihood



- Exotic top: jet **expected to originate from u/c**
  Assume u and c jets equally likely:
  $\epsilon_{u/c} \equiv \frac{1}{2}\left[P(\text{obs. tag} \mid \text{u-jet}) + P(\text{obs. tag} \mid \text{c-jet})\right]$
  e.g. for $\boxed{\text{WP 90 \%:}}$
    - If observed tag is b: $\epsilon_{u/c} = \boxed{0.182}$
    - If observed tag is not b: $\epsilon_{u/c} = 0.818$

average

efficiency operating point	b-efficiency	c-efficiency	light-flavour efficiency
WP 65 %	0.66	0.00944	0.00026
WP 70 %	0.71	0.02081	0.00060
WP 77 %	0.78	0.05870	0.00218
WP 85 %	0.85	0.17505	0.01038
WP 90 %	0.90	0.33187	0.03237

$$-\mathrm{LLH} = \frac{1}{2}\left(\frac{m_{t,aj} - \mu_{t,e}}{\sigma_e}\right)^2 + \frac{1}{2}\left(\frac{m_{t,Wj'} - \mu_{t,l}}{\sigma_l}\right)^2 \boxed{-\log \epsilon_{u/c}} - \log \epsilon_b$$

The ATLAS Collaboration, Transforming jet flavour tagging at ATLAS. Nature Commun. 17 (2026) 541. http://dx.doi.org/10.1038/s41467-025-65059-6

Optimizing the Selection with LLH

Compare

- **Signal: combined samples covering** $m_a = 1, 5, 10$ **GeV and** $c\tau_a = 1, 10, 100$ **mm with BR(**$t \to aq) = 4.6 \cdot 10^{-6}$
- **Three main background processes**

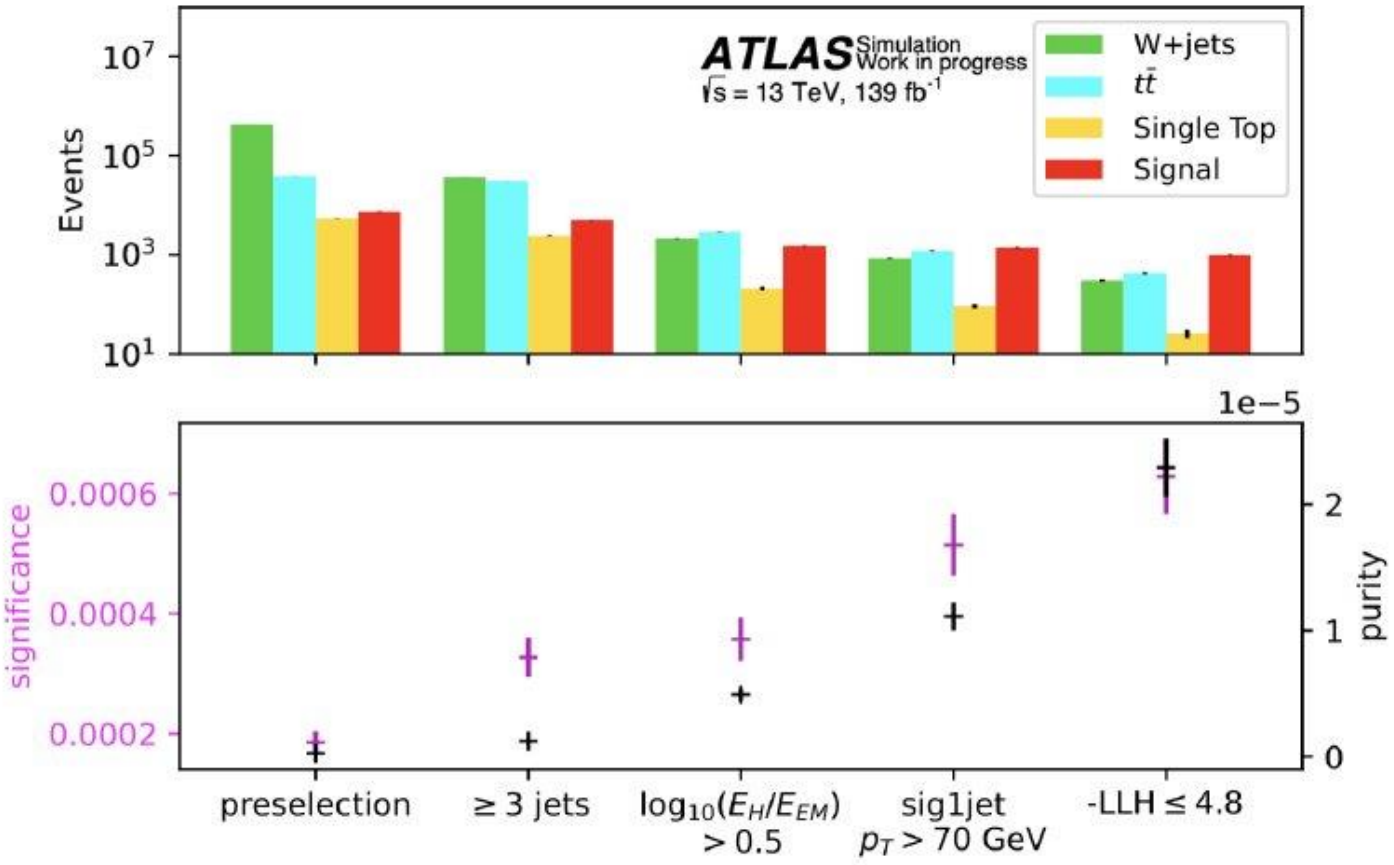
LLH in cutflow optimization:

- **Suppresses background more strongly than signal**

  => higher **purity** $P \approx \frac{S}{B}$

- **Improves statistical sensitivity**

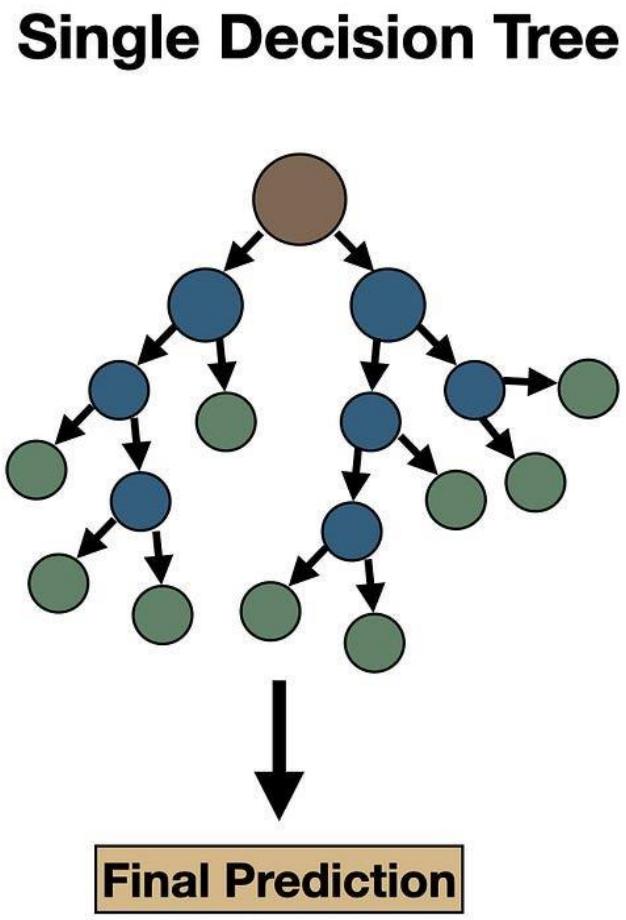
  => higher **significance** $Z \approx \frac{S}{\sqrt{B}}$



Boosted Decision Trees

aim: multivariate signal-background discrimination

- Conventional Cutflow: a fixed sequence of cuts is applied to every event
- Decision tree: a branching sequence of cuts on input variables used for classification

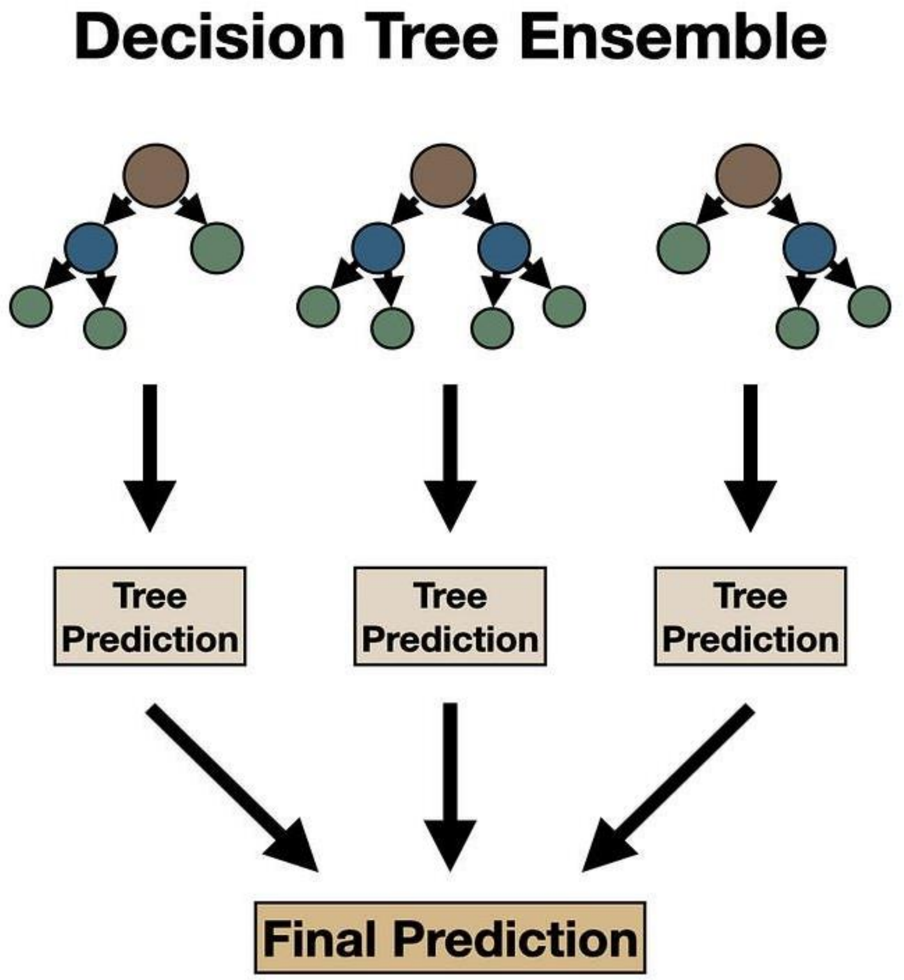


https://rforpoliticalscience.com/wp-content/uploads/2024/03/1_onhgor9hp5rsuz40huzu1a.jpg

Boosted Decision Trees

aim: multivariate signal-background discrimination

- Conventional Cutflow: a fixed sequence of cuts is applied to every event
- Decision tree: a branching sequence of cuts on input variables used for classification
- Boosted: multiple weak trees are trained sequentially and combined into a strong classifier
  1. Start with a simple tree
  2. Evaluate the current ensemble
  3. Add a new tree to correct the ensemble's mistakes and improve the prediction
  4. Combine all trees into a final classifier score



https://rforpoliticalscience.com/wp-content/uploads/2024/03/1_onhgor9hp5rsuz40huzu1a.jpg

ABCD Method

Challenges with MC simulations:

- Poor statistics in the signal-like region
- QCD multijet background is difficult to model reliably
- Detector mismodelling can bias background predictions

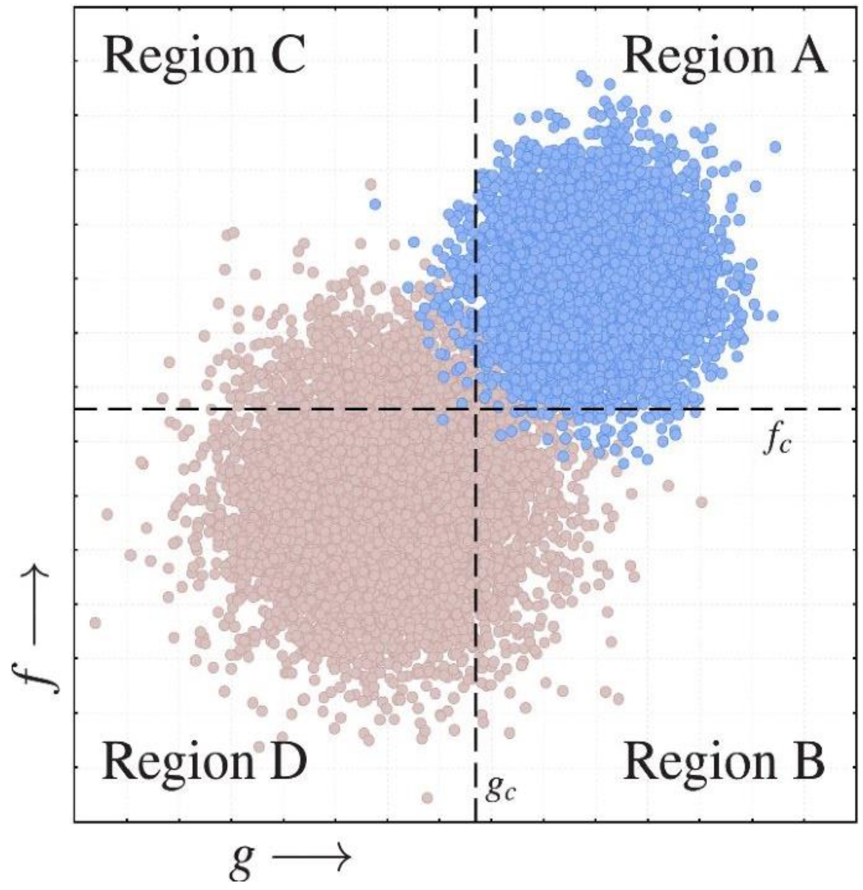
We need two parameters, f and g, that:

- Discriminate signal from background => BDT scores
- Are approximately uncorrelated for background

Then predict background in A from B/C/D: $N_A^{\text{pred}} = \frac{N_B \cdot N_C}{N_D}$

Validate the prediction on MC with a Closure Test: $\frac{N_A^{\text{obs}}}{N_A^{\text{pred}}} \stackrel{?}{=} 1$

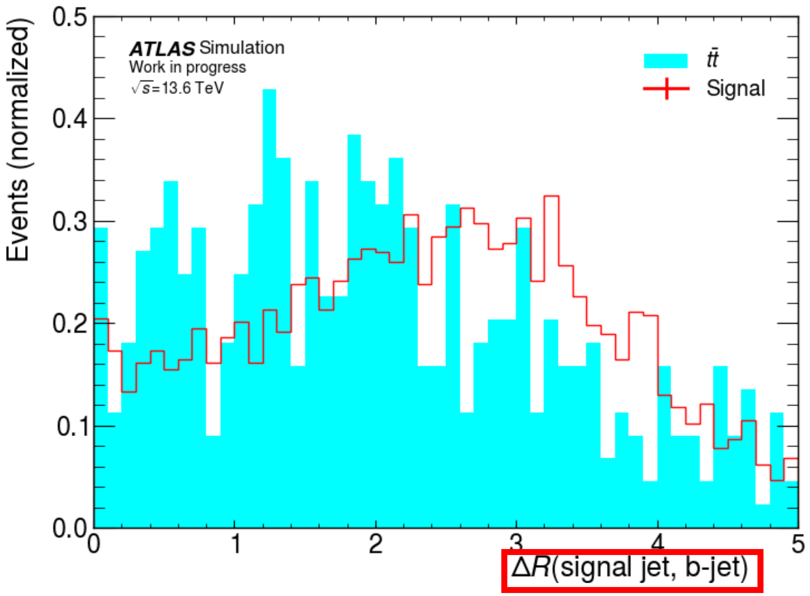
Deviation from 1 => non-closure => systematic uncertainty

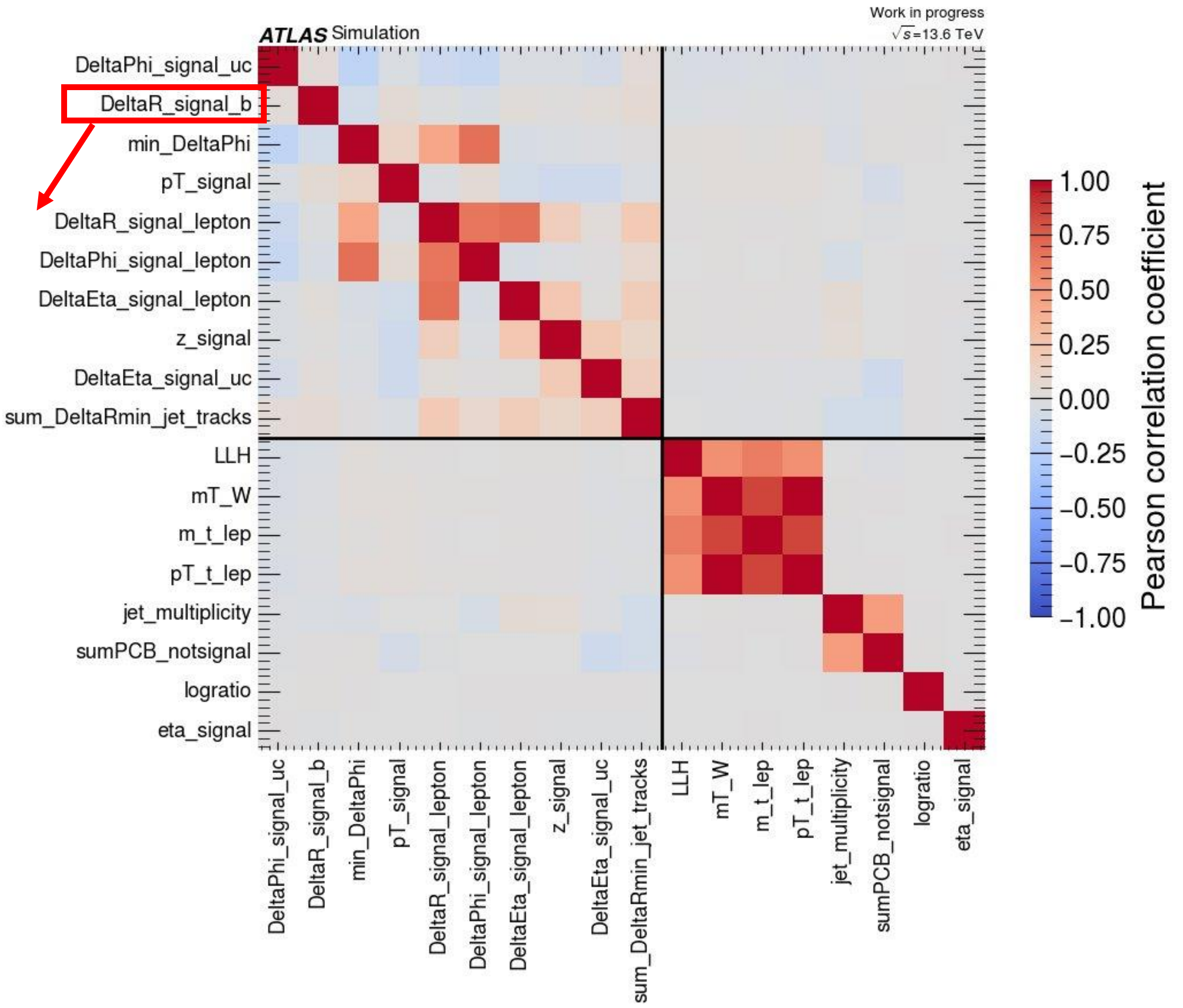


Bardhan, J., Mandal, T., Mitra, S. *et al.* Unsupervised and lightly supervised learning in particle physics. *Eur. Phys. J. Spec. Top.* **233**, 2559–2596 (2024). https://doi.org/10.1140/epjs/s11734-024-01235-x

Correlation Matrix

- Many observables can potentially discriminate signal from background
- Strategy:
  - Combine observables into two separate BDTs
  - Ensure that the BDT outputs are largely uncorrelated for background

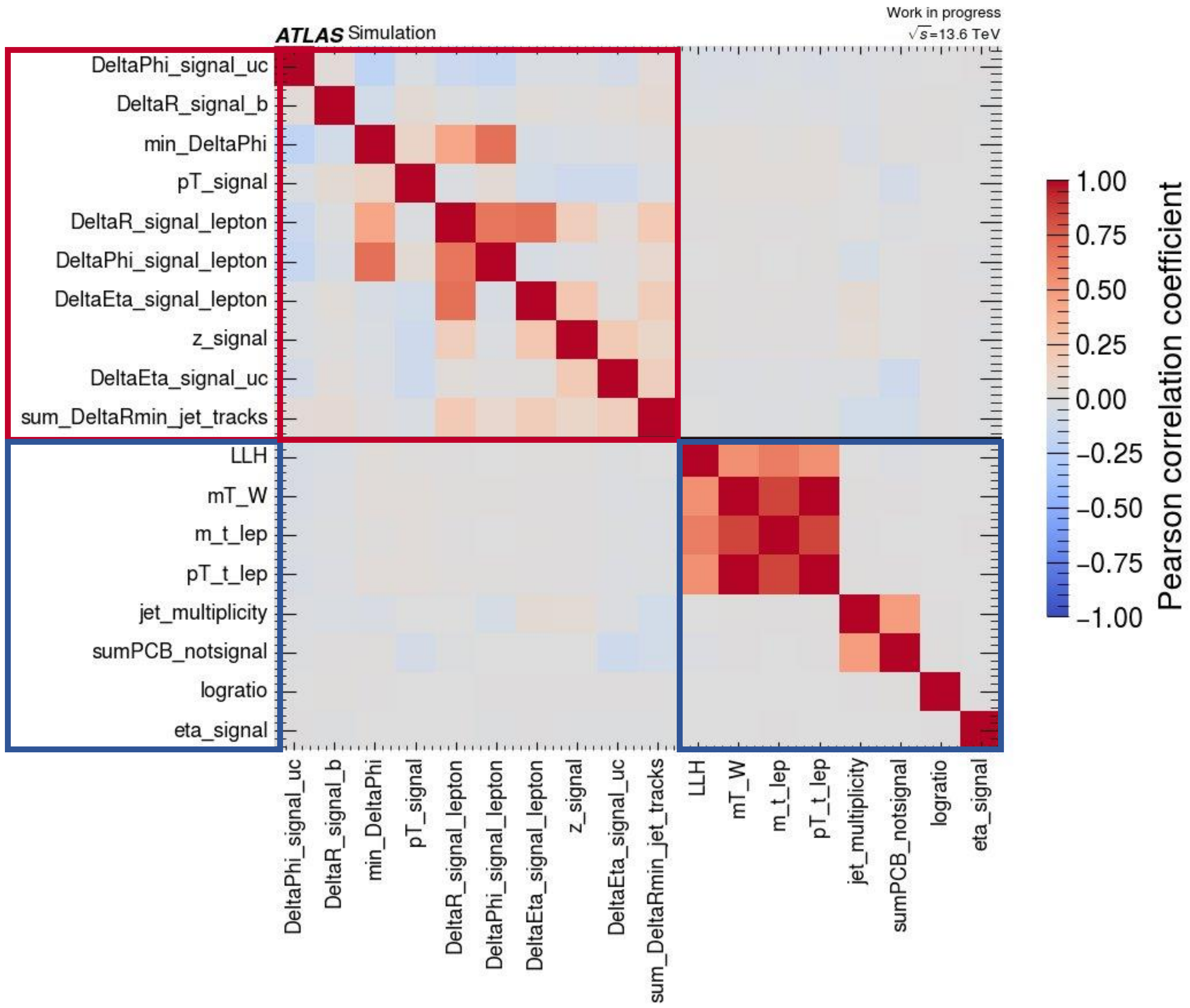




Decorrelation Strategy

Perfect decorrelation of the BDT inputs is not achievable, therefore:

- **BDT 1: Gradient Boosting (*XGBoost*)**
  - **Calculate residuals and define a loss function**
  - **Fit each new tree to the remaining prediction error (negative gradient)**
- **BDT 2: Uniform Gradient Boosting**
  - **Uniform: BDT 2 response for background largely independent of BDT 1 score**
  - **Modified loss function penalizes non-uniformity**

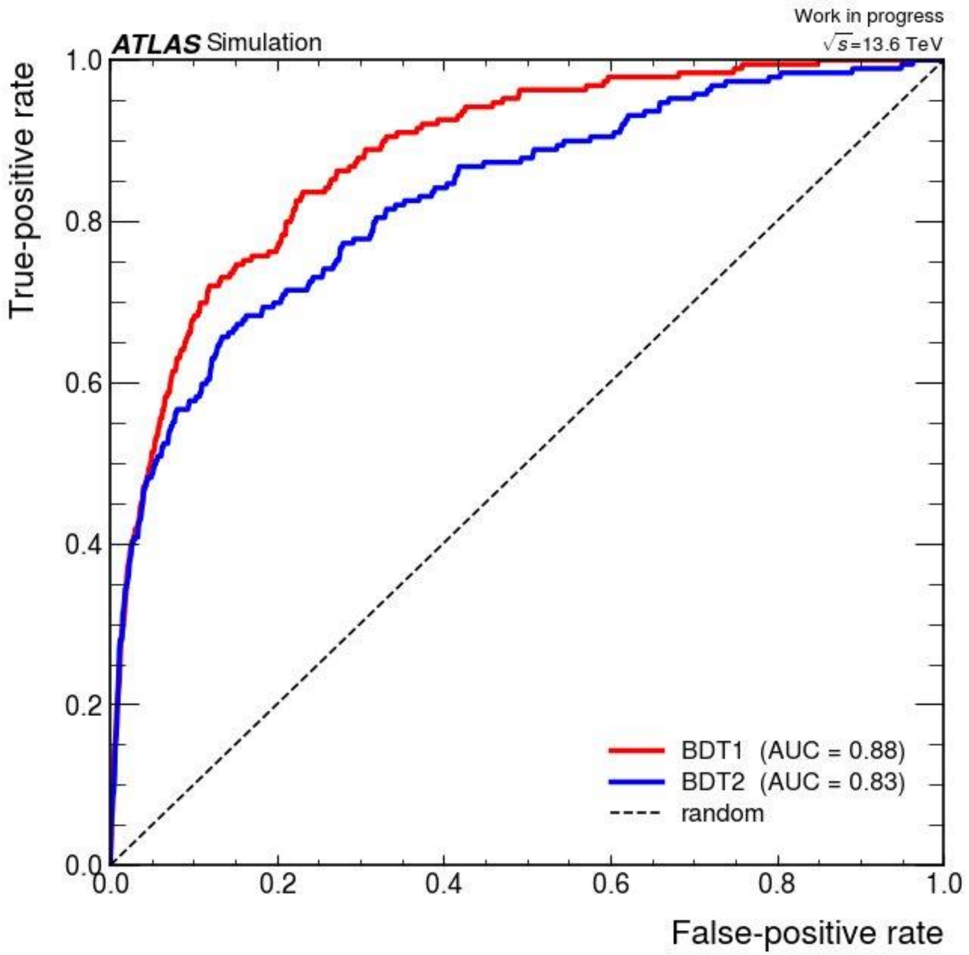


Receiver Operating Characteristic

ROC Curve:

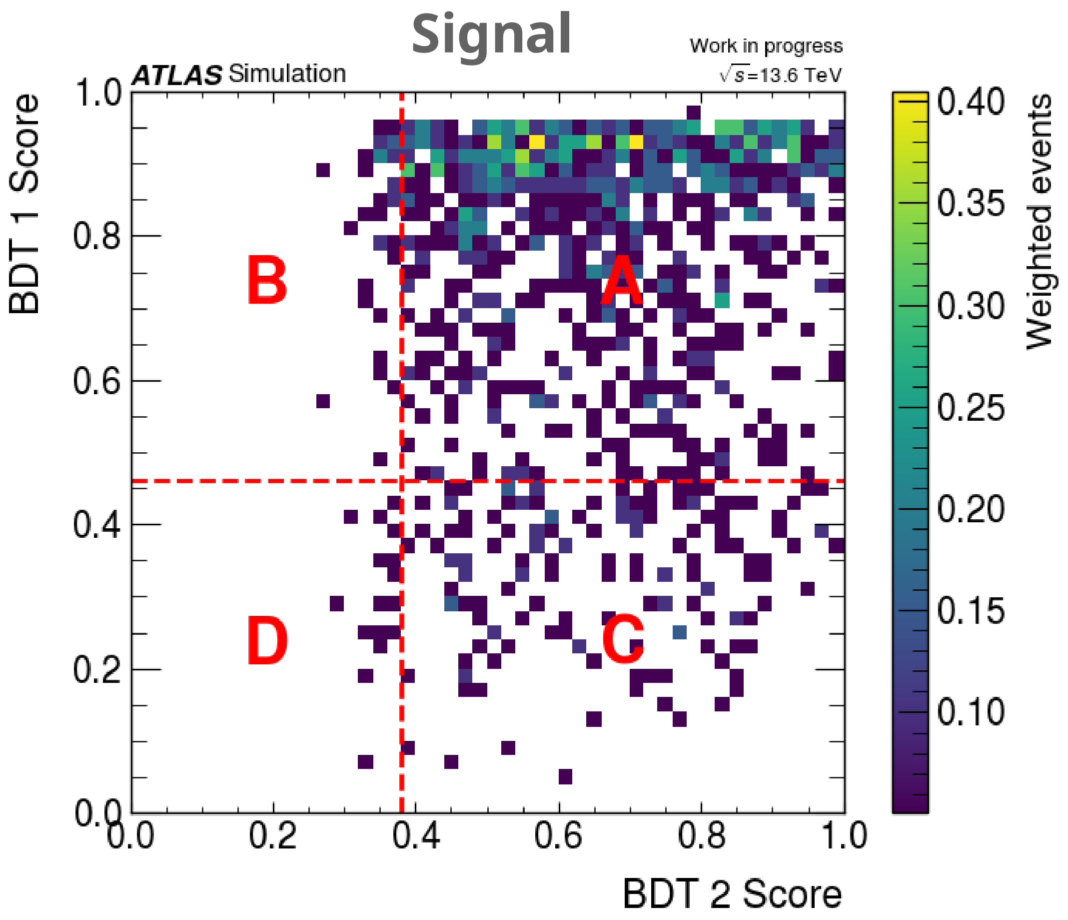
- **Each point corresponds to a different cut on the BDT score**
- **Shows the trade-off between signal efficiency (TPR) and background rejection (1 – FPR)**
- **Better classifiers lie closer to the upper-left corner**
- **BDT2 retains good performance despite decorrelation**

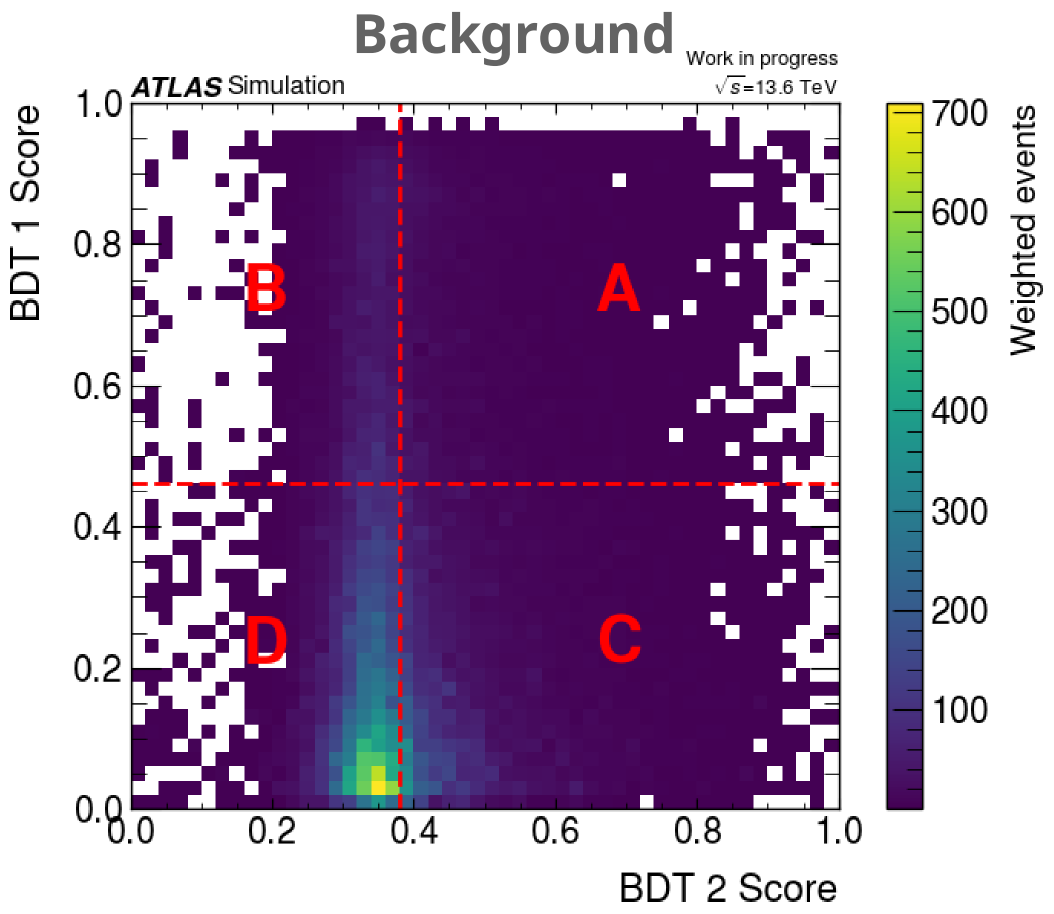
Background Sample	BDT1 AUC	BDT2 AUC
combined BG	0.88	0.83
W+jets	0.89	0.86
ttbar	0.84	0.73
single top	0.84	0.70



BDT Score Distributions

- Signal highly concentrated in region A
- Background broadly populates the control regions, with only a small contribution in region A

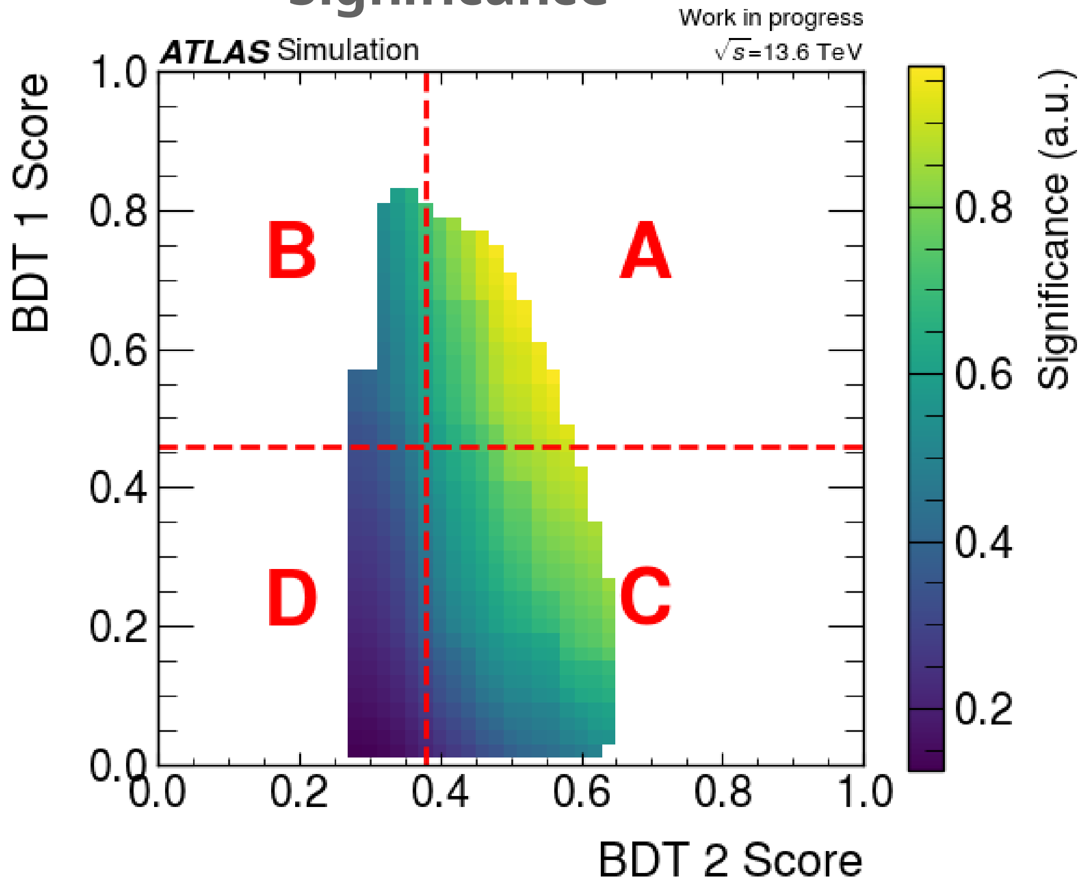




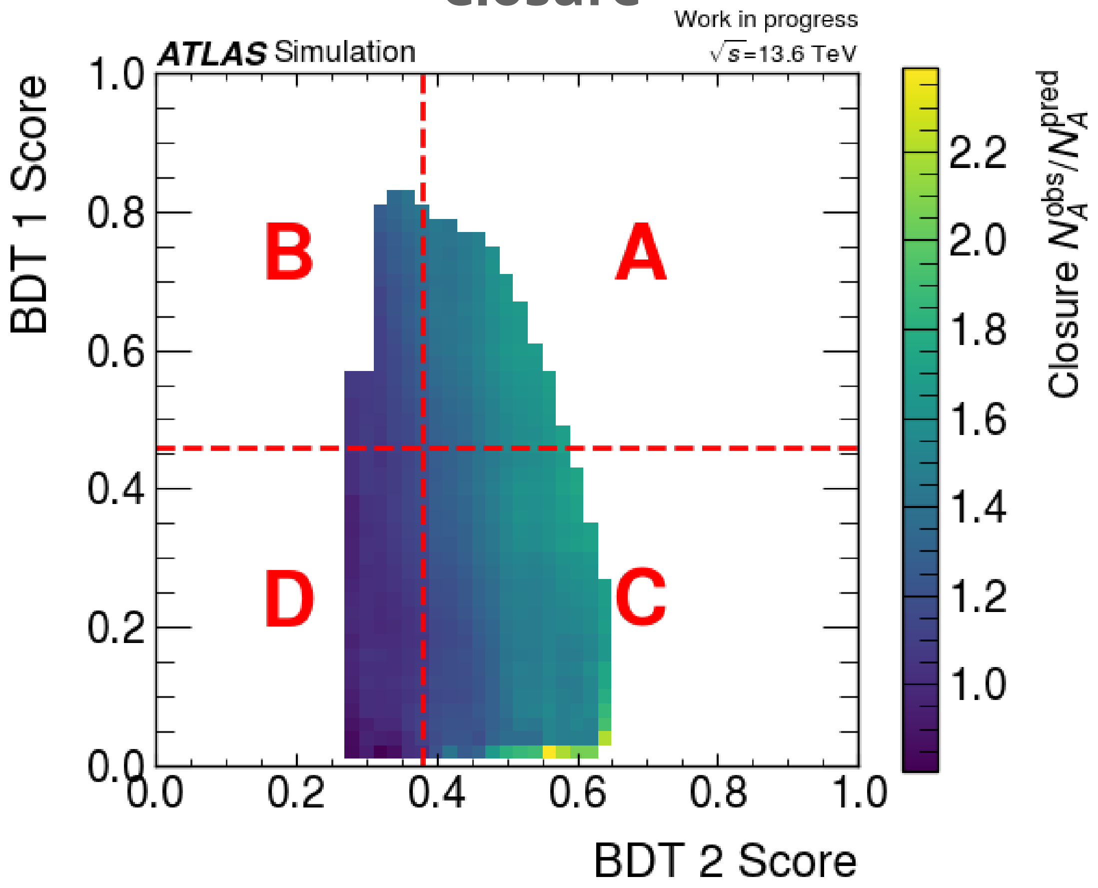
Optimizing ABCD Regions

Best significance and best closure occur in different regions of BDT space

Significance



Closure



Optimizing ABCD Regions

- Intersection point chosen
  - To maximize significance
  - Subject to a constraint: closure within 20% of unity
- Position roughly central in BDT space ensures:
  - Balanced use of both BDT scores
  - Stable background prediction
  - Good statistics in all ABCD regions
- In numbers:
  - Observed BG in signal region: $N_{A,obs}$ = 3450 ± 45
  - Prediction from control regions: $N_{A,pred}$ = 2910 ± 70
  - Closure: 1.18 ± 0.03

=> Reasonable working point for the current setup; further optimization of the analysis required

Conclusions

Summary:

- Ongoing analysis of the charming ALP model with ATLAS
- Aim: constrain the flavor-changing **top-ALP coupling** for ALP masses in the GeV range
- Jet assignment and event selection using **LogRatio and LLH**
- Data-driven background estimation using **BDTs** within the **ABCD method**

Outlook:

- Process additional simulated events to improve statistics
- Refine the analysis strategy and apply it to collision data
- Obtain an **upper limit on the branching ratio**

Backup Slides

BDT Parameters

```
return XGBClassifier(
    objective="binary:logistic",
    n_estimators=500,
    max_depth=3,
    learning_rate=0.02,
    verbosity=0,
    gamma=0.5,
    min_child_weight=10,
    subsample=0.7,
    colsample_bytree=0.8,
    scale_pos_weight=1,
    base_score=0.5,
    eval_metric=["auc", "logloss"],
    n_jobs=4,
)
```

```
loss = KnnFlatnessLossFunction(
uniform_features=["bdt1_score"],
uniform_label=0,
n_neighbours=700,
fl_coefficient=50,
)

bdt2 = UGradientBoostingClassifier(
loss=loss,
n_estimators=150,
max_depth=5,
min_samples_leaf=40,
learning_rate=0.05,
subsample=0.8,
train_features=list(variables_kin),
)
```

Constraints

Constraint type	Requirement
Signal retention in A	N_A^sig > 0.5 × N_total^sig
Region A background	N_A^bkg >= 2.0
Region B background	N_B^bkg >= 20.0
Region C background	N_C^bkg >= 20.0
Region D background	N_D^bkg >= 50.0
Region A signal	N_A^sig >= 1.0
Region B signal	N_B^sig >= 0.1
Region C signal	N_C^sig >= 0.0
Closure lower bound	closure >= 0.8
Closure upper bound	closure <= 1.2
Closure finite	closure must be finite
Closure uncertainty finite	closure_err must be finite

BDT Variables

Variable	Description
$p_T^{sig1jet}$	Transverse momentum component of the signal candidate jet
sig1jet QCD score	NN QCD score of the signal candidate jet
$\Delta\phi$(sig1jet, lepton)	Opening angle between the signal candidate jet and the lepton
$\Delta\eta$(sig1jet, lepton)	Opening distance between the signal candidate jet and the lepton
min($\Delta\phi$(sig1jet, lepton), $\Delta\phi$(sig1jet, E_T^{miss}))	Minimum opening angle between the signal candidate and the lepton and the missing transverse energy
p_T^{lepton}	Transverse momentum component of the lepton
ΔR(lepton, closest jet)	Angular distance between the lepton and the closest clean jet
E_T^{miss}	Missing transverse energy
$m_{T,W}$	W candidate transverse mass
m_{inv}(lepton, closest jet)	Invariant mass of the lepton and the closest clean jet
m_{inv}(lepton, sig1jet)	Invariant mass of the lepton and the signal candidate jet
m_{inv}(all jets, not sig1jet)	Invariant mass of all clean jets in the event excluding the signal candidate jet
mean NN signal clean jets	Mean NN signal score value of all clean jets in the event
$\Delta\phi$(sig1jet, u/c-jet)	Opening angle between the signal candidate jet and the reconstructed u/c-jet candidate
$\Delta\eta$(sig1jet, u/c-jet)	Opening distance between the signal candidate jet and the reconstructed u/c-jet candidate
ΔR(sig1jet, u/c-jet)	Angular distance between the signal candidate jet and the reconstructed u/c-jet candidate
$\Delta R(u/c\text{-jet}, b\text{-jet})$	Angular distance between the reconstructed u/c-jet candidate and the reconstructed b-jet candidate

BDT Variables

$m_{inv}(\text{sig1jet}, W)$	Invariant mass of the signal candidate jet and the W boson candidate
$\Delta R(\text{sig1jet, lepton})$	Angular distance between the signal candidate jet and the lepton
$\Delta R(u/c\text{-jet, lepton})$	Angular distance between the reconstructed u/c-jet candidate and the lepton
$\Delta R(\text{sig1jet}, b\text{-jet})$	Angular distance between the signal candidate jet and the reconstructed b-jet candidate
p_T^{tlep}	Transverse momentum component of the reconstructed leptonically decaying Top quark
p_T^{texot}	Transverse momentum component of the reconstructed exotically decaying Top quark
$\sum$ PCB, no sig1jet	Sum of the PCB scores of all jets in the event excluding the signal candidate jet
sig1jet signal score	NN signal score of the signal candidate jet
$L_{xy}^{sig1jet}$	Jet cluster showering start position in the $x-y$-plane of the signal candidate jet
$z^{sig1jet}$	Jet cluster showering start z-position of the signal candidate jet
$\eta^{sig1jet}$	Pseudorapidity of the signal candidate jet
sig1jet logRatio	$\log_{10}(E_H/E_{EM})$ of the signal candidate jet
jet multiplicity	Number of jets in the event
$m_{inv}(t_{lep})$	Invariant mass of the reconstructed Top quark that decays leptonically
$m_{inv}(t_{exot})$	Invariant mass of the reconstructed Top quark that decays exotically
-LLH	Negative loglikelihood