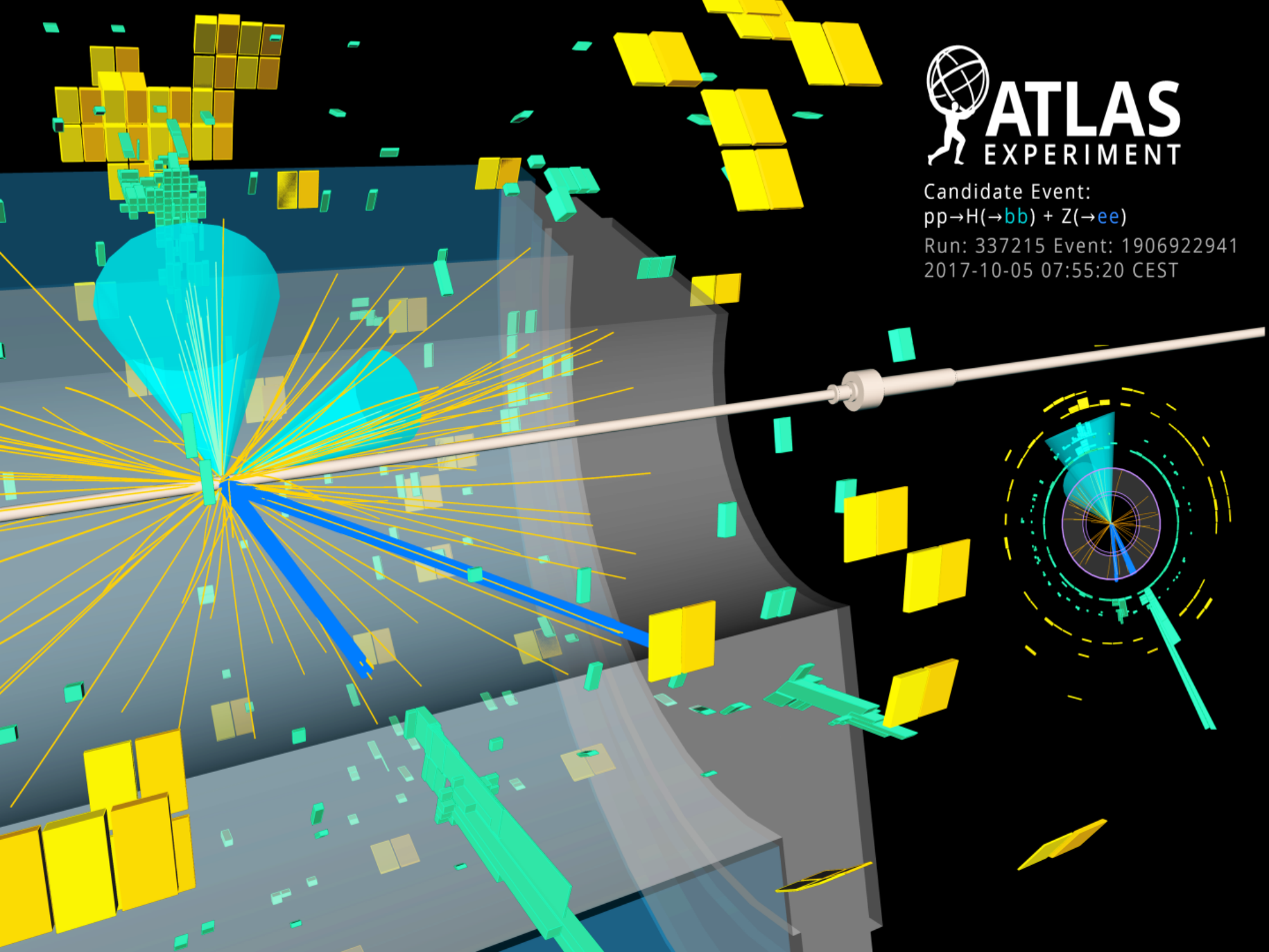
Charm tagging and search for $ZH \to llc\bar{c}$ decay with ATLAS data

SUPRIYA SINHA

Supervisor: Dr. Tatjana Lenz

Prof. Jochen Dingfelder

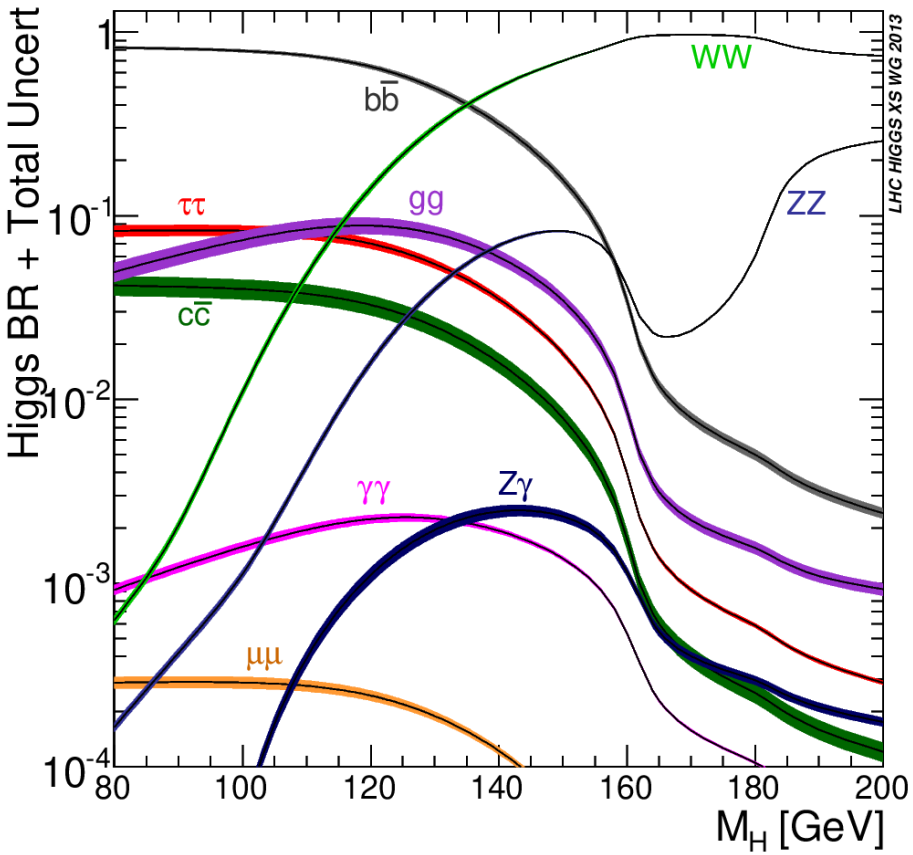
27 August, 2019


ATLAS
EXPERIMENT
Candidate Event:
pp→H(→bb) + Z(→ee)
Run: 337215 Event: 1906922941
2017-10-05 07:55:20 CEST

Outline

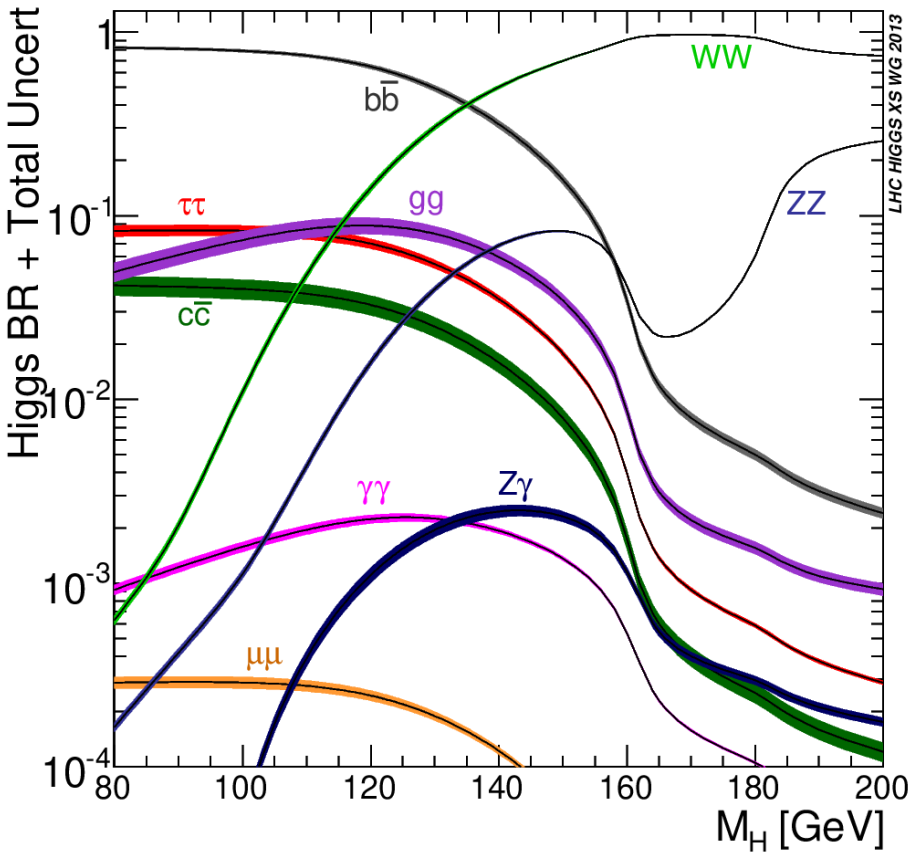
Moving on

Higgs decay



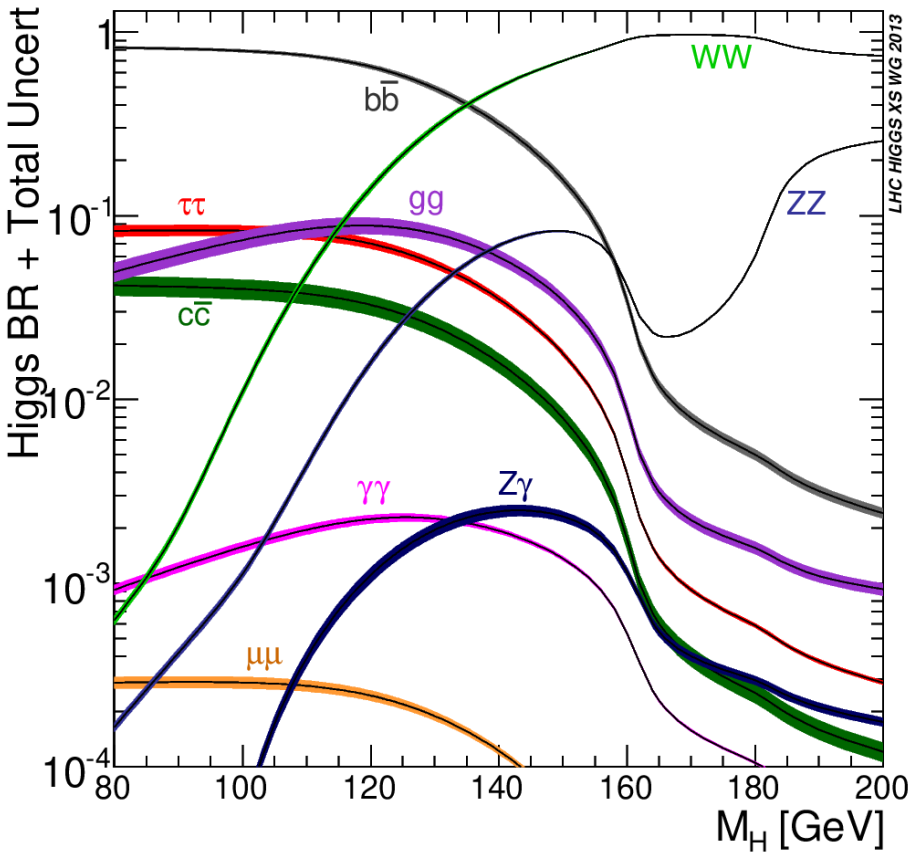
- Fermionic: $b\bar{b}, \tau\tau, c\bar{c}$, etc.; Bosonic: $WW, ZZ, \gamma\gamma$, etc.

Higgs decay



- Fermionic: $b\bar{b}, \tau\tau, c\bar{c}$, etc.; Bosonic: $WW, ZZ, \gamma\gamma$, etc.
- Decays not just to third generation particles, but second generation particles too, like $c\bar{c}$

Higgs decay

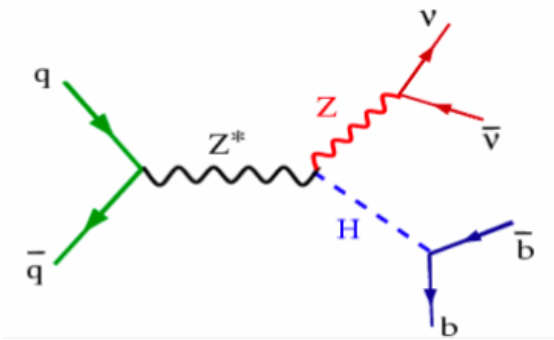


- Fermionic: $b\bar{b}, \tau\tau, c\bar{c}$, etc.; Bosonic: $WW, ZZ, \gamma\gamma$, etc.
- Decays not just to third generation particles, but second generation particles too, like $c\bar{c}$
- $q\bar{q} \rightarrow ZH \rightarrow l\bar{l}c\bar{c}$ channel taken for analysis

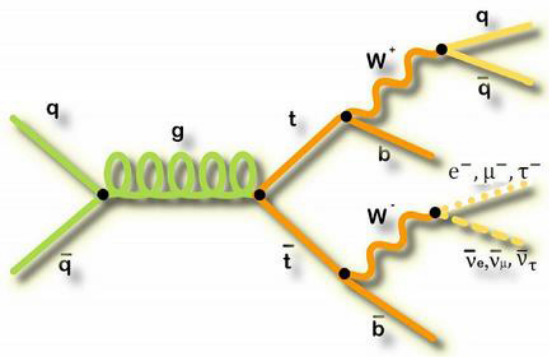
Moving on

b- tagging

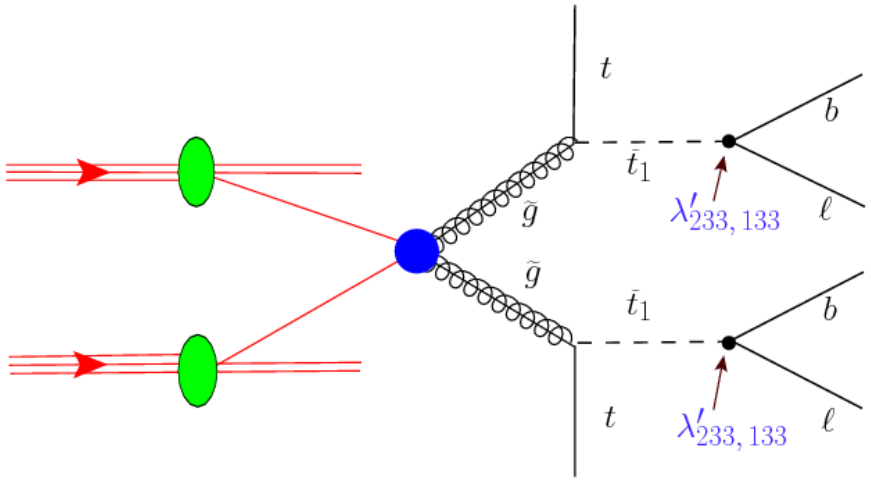
- Identification of B-hadrons in a jet
- Often used in analyses containing high p_T b-jets in their final states



Higgs decay to bottom quark [1]



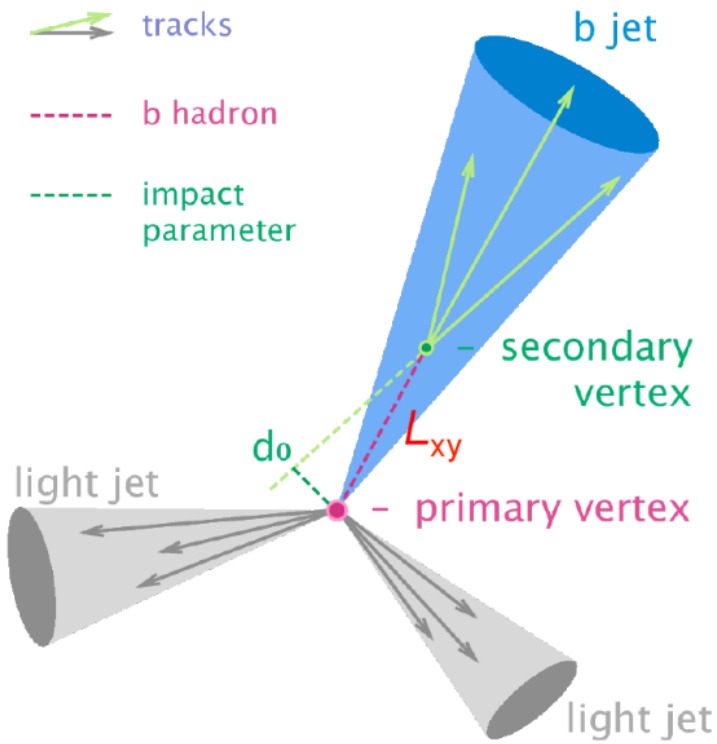
Top physics [2]



BSM searches [3]

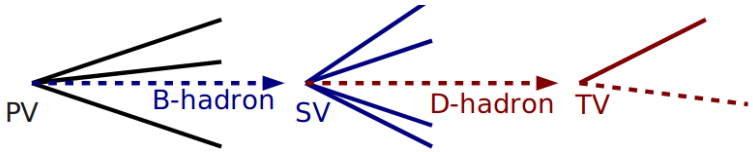
B-hadron properties

- Large momentum fraction of b-hadrons from b-quark hadronization



Secondary vertex displaced from the primary vertex [4]

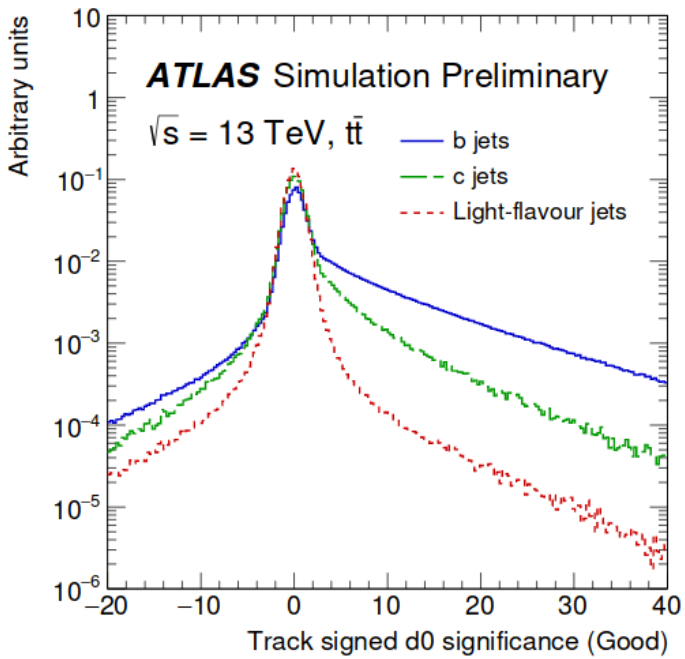
- Long lifetime of b-hadrons (1.5 ps) $\Rightarrow$ Secondary Vertex
- High mass of b-quarks $\Rightarrow$ large p_T product $\Rightarrow$ broad angular distribution
- B-hadron decay product $\Rightarrow$ Tertiary vertex



Tertiary vertex from D-hadron decay [5]

b-tagging algorithms

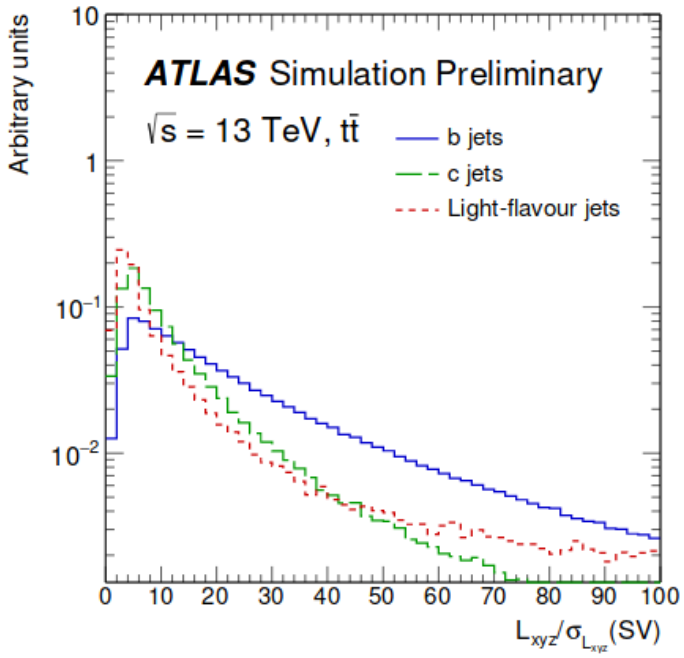
- Search for a variable sensitive to the flavor content of the jet $\Rightarrow$ used as a discriminator
- b-tagging methods:
  - Impact parameter based: SV tracks have large IP w.r.t. the PV
    Define IP significance: $s_{d_0} = \frac{d_0}{\sigma_{d_0}}$



[6]

b-tagging algorithms

- Search for a variable sensitive to the flavor content of the jet $\Rightarrow$ used as a discriminator
- b-tagging methods:
  - Secondary vertex based: SV candidates are reconstructed to get SV point

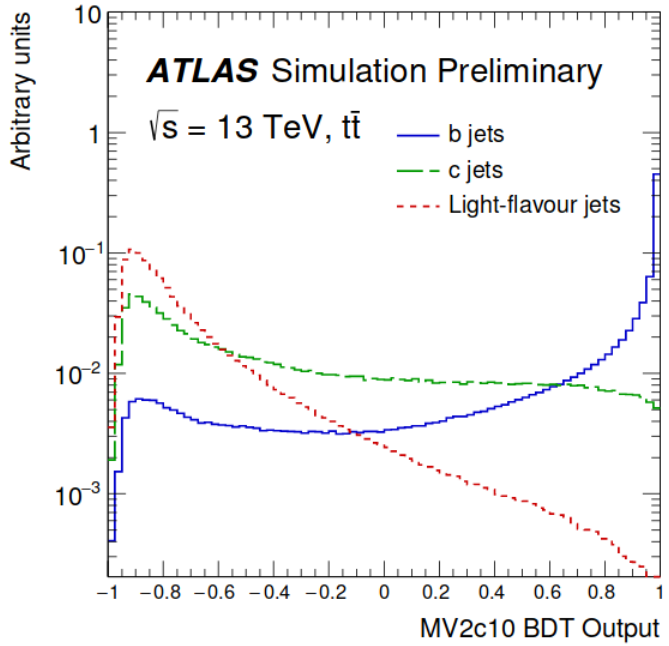


[6]

Discrimination based on properties of reconstructed vertex: L_{XY}, invariant mass of the SV, etc.

b-tagging algorithms

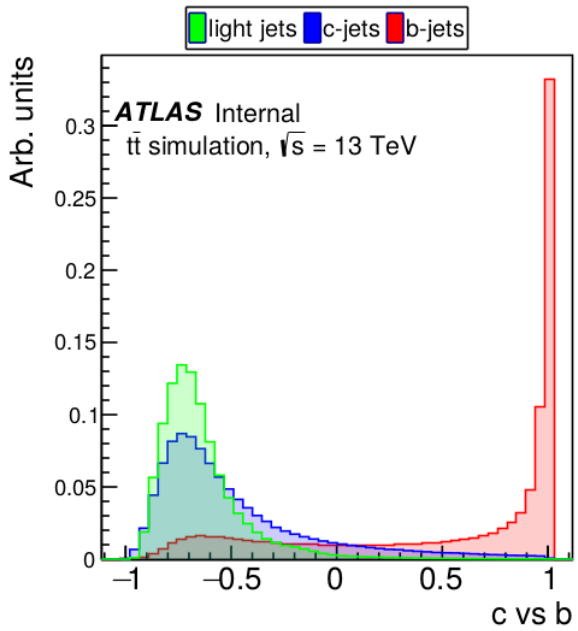
- Search for a variable sensitive to the flavor content of the jet ⇒ used as a discriminator
- b-tagging methods:
  - Combined algorithms: Combines different variables
    Usage of Boosted Decision Trees (BDTs), Neural Networks (NN), etc.

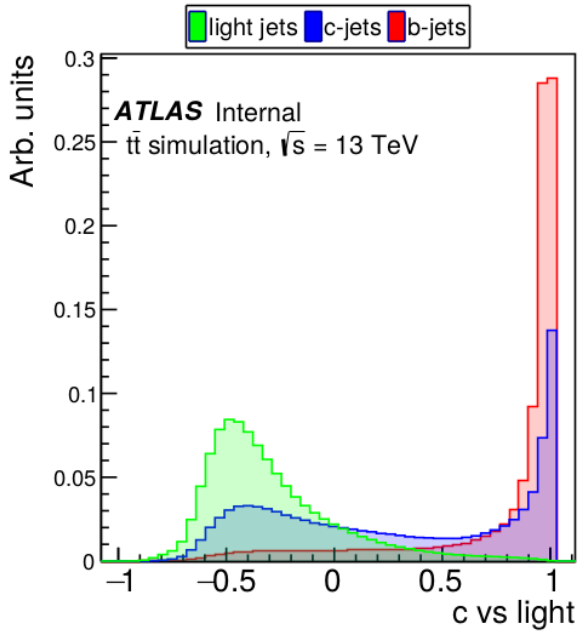


[6]

c-tagging

- Similar methods as b-tagging
- Relevant c-hadron properties:
  - Relatively long lifetime: $0.5 - 1$ *ps* (still shorter than b-hadrons)
  - Decay to a smaller number of charged particles as compared to b-hadrons
- Two multivariate discriminants are trained: separating c-jets from l-jets; separating c-jets from b-jets





[7]

c-tagging

- Similar methods as b-tagging
- Relevant c-hadron properties:
  - Relatively long lifetime: $0.5 - 1$ *ps* (still shorter than b-hadrons)
  - Decay to a smaller number of charged particles as compared to b-hadrons
- Two multivariate discriminants are trained: separating c-jets from l-jets; separating c-jets from b-jets

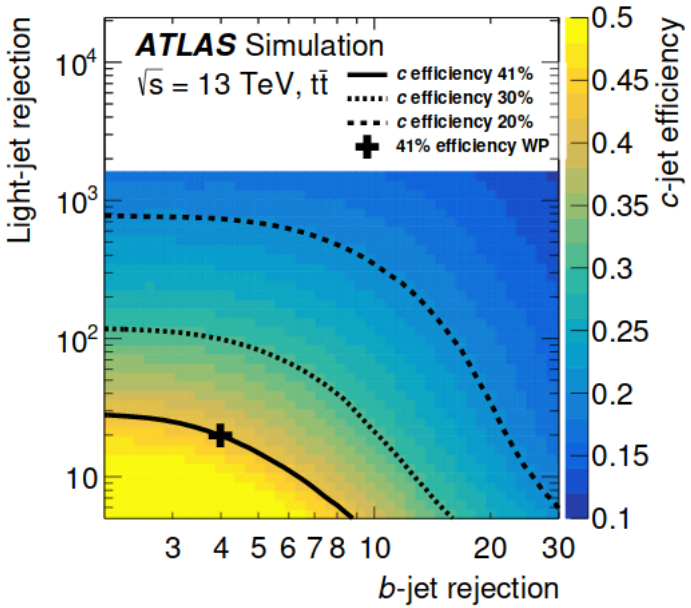


Fig: c-jet tagging efficiency as a function of b-jet and l-jet rejection [7]

Moving on

c-tagging using D^* reconstruction

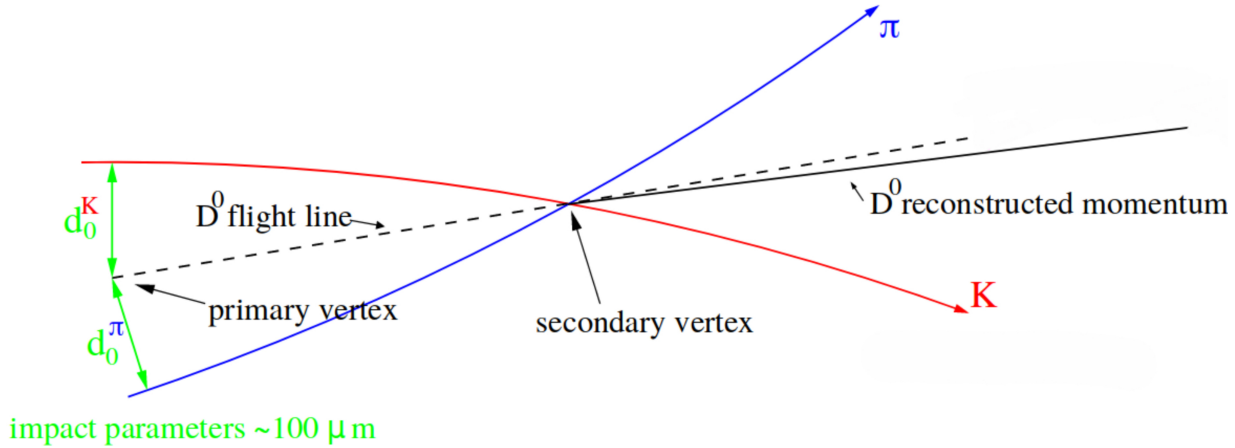
- Require: a channel with a strong predominance of charm flavour
- c- quark hadronization:

$$f(c \to D^0) = 0.557 \pm 0.023,$$
$$f(c \to D^+) = 0.226 \pm 0.010,$$
$$f(c \to D^{*+}) = 0.238 \pm 0.007.$$

- Reconstructing: $D^{*+} \to D^0\pi^+ \to (K^-\pi^+)\pi^+$

  BR $(D^{*+} \to D^0\pi^+) \approx 67.7\%$

  BR $(D^0 \to K^-\pi^+) \approx 3.9\%$

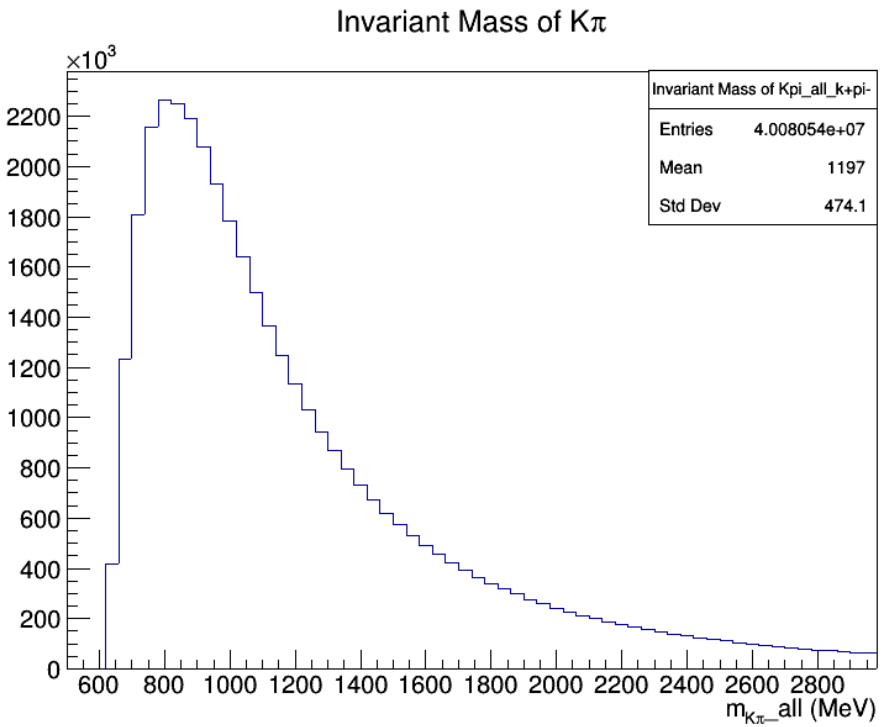


c-tagging using D^* reconstruction

- Need: a channel with strong predominance of charm flavour
- Reconstructing: $D^{*+} \rightarrow D^0\pi^+ \rightarrow (K^-\pi^+)\pi^+$
- Selection and reconstruction:
  - $trkP_T > 1\ GeV$
  - Kaon and Pion mass hypothesis for each track
  - Considered all possible combinations of oppositely charged $K\pi$ candidates per jet

c-tagging using D^* reconstruction

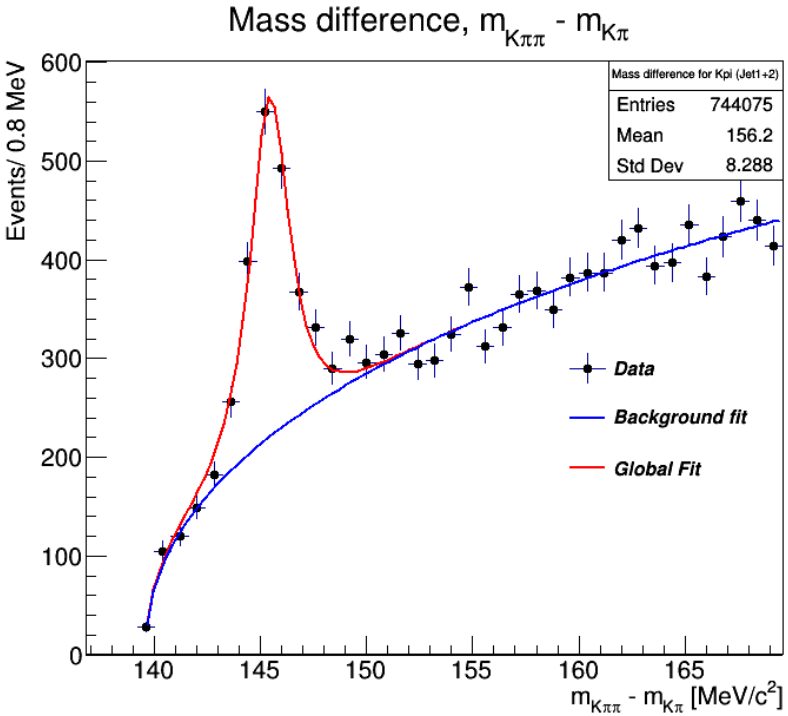
- Need: a channel with strong predominance of charm flavour
- Reconstructing: $D^{*+} \rightarrow D^0\pi^+ \rightarrow (K^-\pi^+)\pi^+$
- Selection and reconstruction:
  - $trkP_T > 1\ GeV$
  - Kaon and Pion mass hypothesis for each track
  - Considered all possible combinations of oppositely charged $K\pi$ candidates per jet

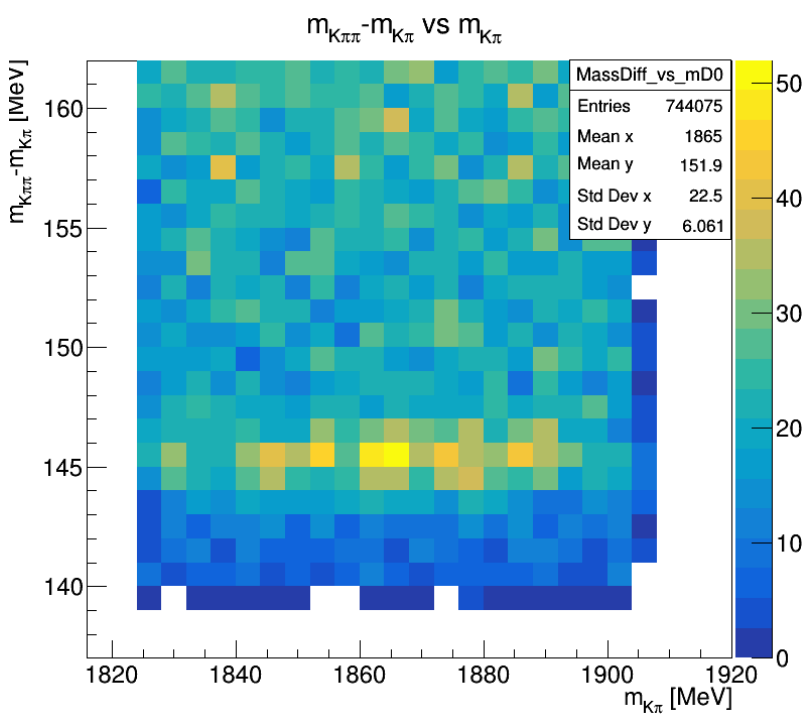


c-tagging using D^* reconstruction

- Need: a channel with strong predominance of charm flavour
- Reconstructing: $D^{*+} \to D^0\pi^+ \to (K^-\pi^+)\pi^+$
- Selection and reconstruction:
  - $trkP_T > 1\ GeV$
  - Kaon and Pion mass hypothesis for each track
  - Considered all possible combinations of oppositely charged $K\pi$ candidates per jet
  - $K\pi$ mass window: $\mid m_{K\pi} - 1864.83 \mid < 40\ MeV$
  - Vertex reconstruction of all candidates within this window
  - For $K\pi\pi$ mass reconstruction:
    - Third track (slow pion) with charge opposite to the corresponding kaon
    - Distance between slow pion track and the reconstructed $K\pi$ vertex $\leq 1\ mm$ in both, the xy-plane and the z-axis
    - $\Delta R(\pi_{slow}, K\pi) \leq 0.15$
    - $p_T(K\pi\pi) > 20\ GeV$
    - $\frac{p_T(K\pi\pi)}{\Sigma p_T} > 0.2$, where Σp_T is the sum of trk p_T in a cone of $\Delta R \leq 0.4$ around the direction of D^* momentum

Reconstructed D^*





- Signal fit (Modified Gaussian):

$$S = Gauss^{mod} \propto exp[-0.5.x^{(1+\frac{1}{1+0.5x})}], \text{where } x = \left|\frac{(\Delta m - \Delta m_0)}{\sigma}\right|$$

Fit values: $\Delta m_0[MeV] = 145.43 \pm 0.06,\ \sigma(\Delta m_0)[MeV] = 0.87 \pm 0.07$

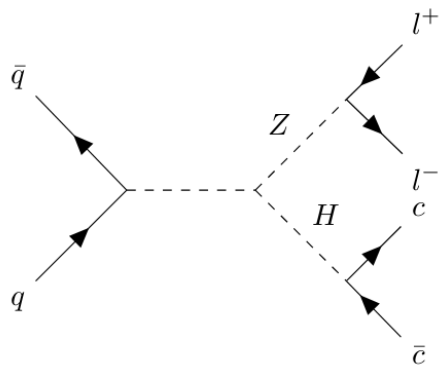
World average: $\Delta m_0 = 145.421 \pm 0.010\ MeV$

- Background fit: $B \propto (\Delta m - m_\pi)^\alpha e^{-\beta(\Delta m - m_\pi)}$

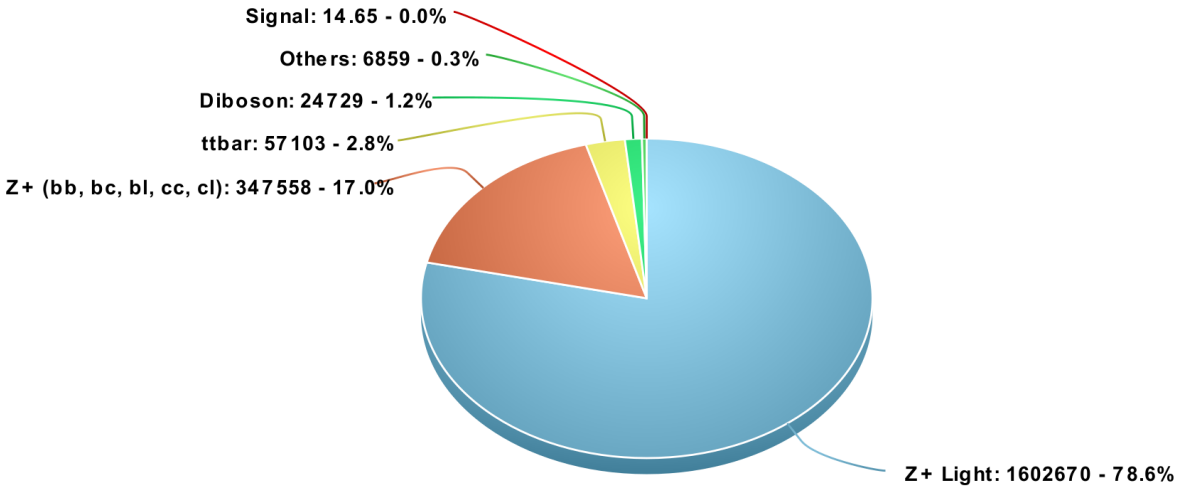
Moving on

Signal and backgrounds

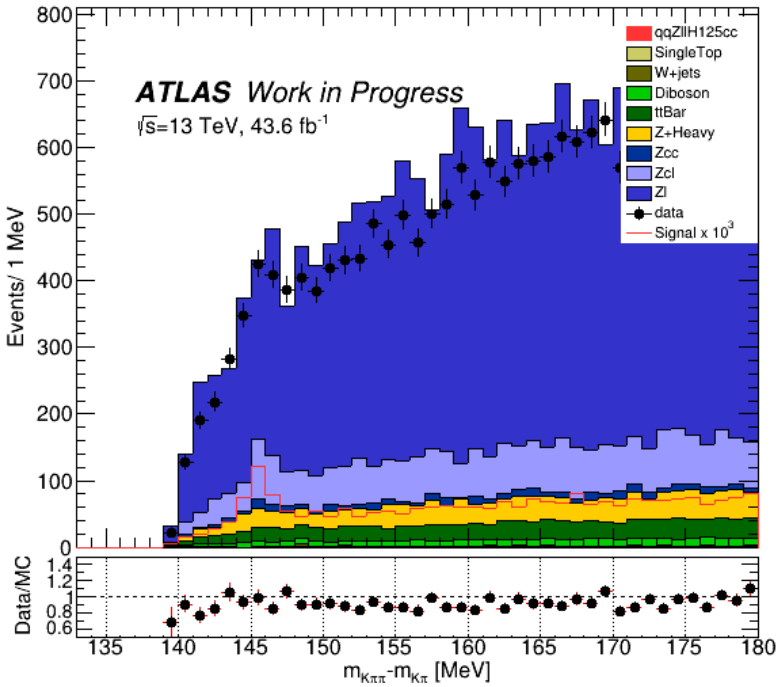
- Signal: $qq \to ZH$, with $ZH \to l^+l^- c\bar{c}$, where $l = e, \mu$

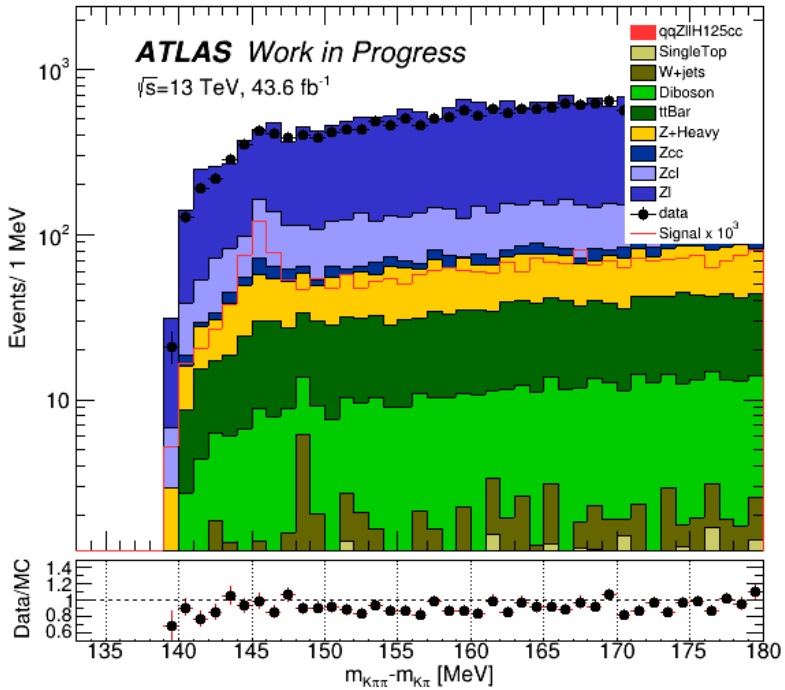


- SM branching fraction for Higgs decay to a pair of charm quarks is predicted to be 2.9%
- Backgrounds: Z+jets, $t\bar{t}$, Diboson (ZZ, WZ, WW) and a small contribution from W+jets and single top



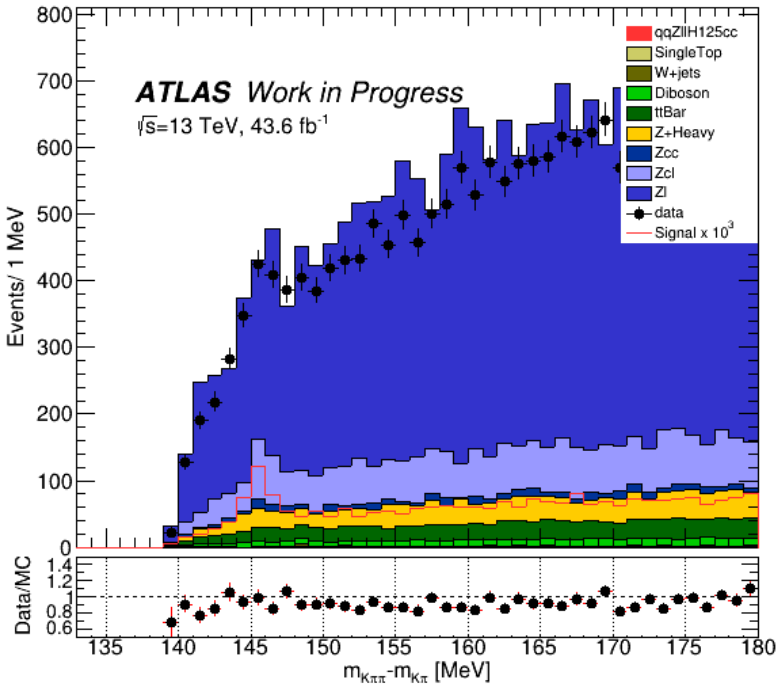
Signal and background

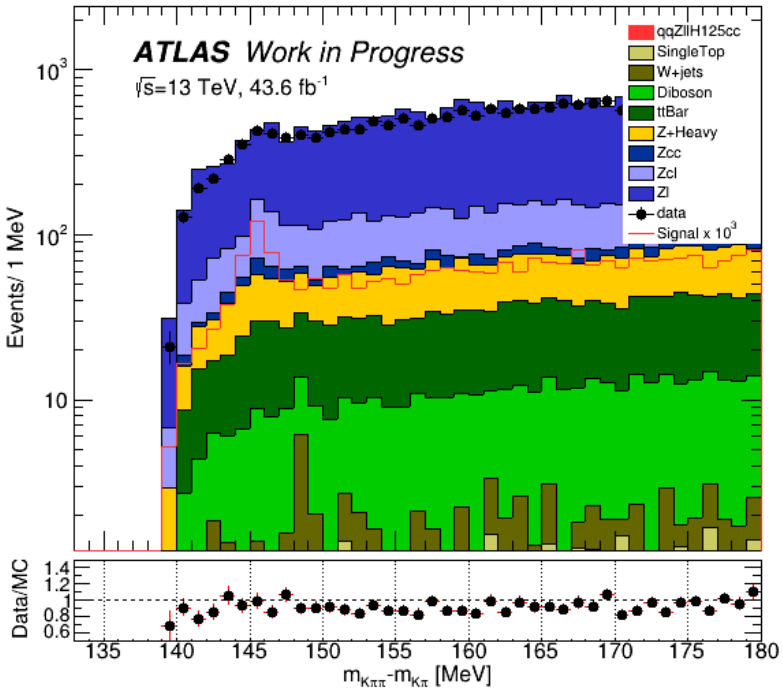




- Major contribution: Z+ light jets
- Signal peaks at 145 *MeV*
- Backgrounds from Zcl and Zcc show a peak too

Signal and background

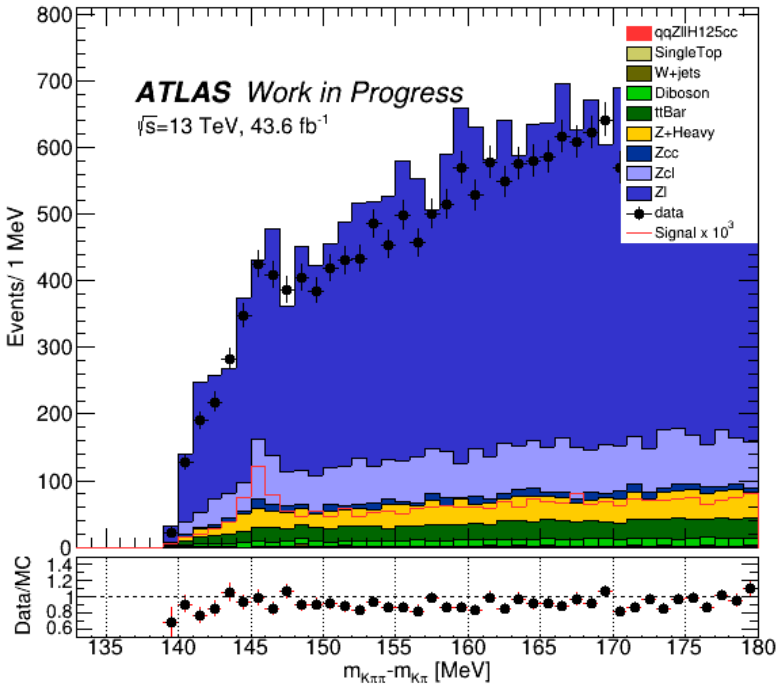


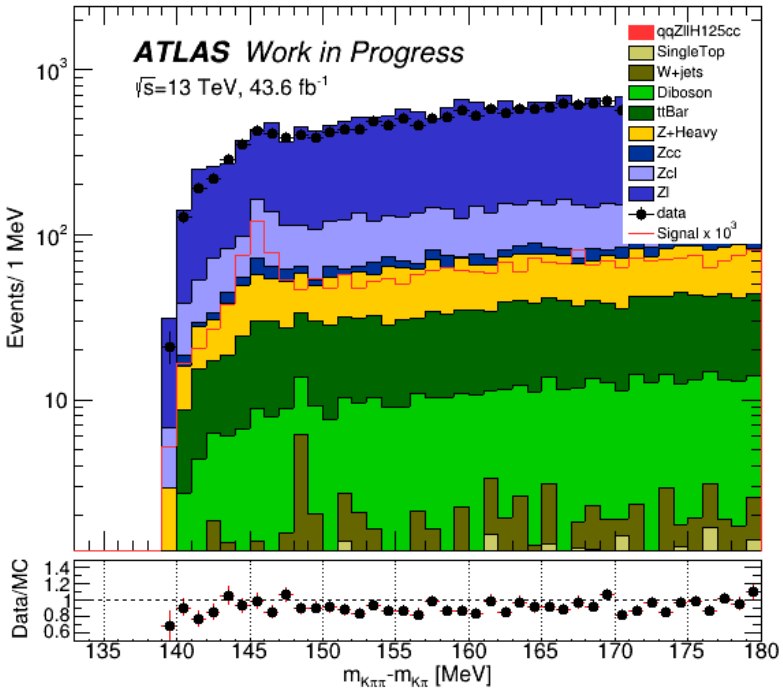


- Major contribution: Z+ light jets
- Signal peaks at 145 *MeV*
- Backgrounds from Zcl and Zcc show a peak too

$\Rightarrow$ Must find a suitable parameter to distinguish the signal from the backgrounds:

Signal and background



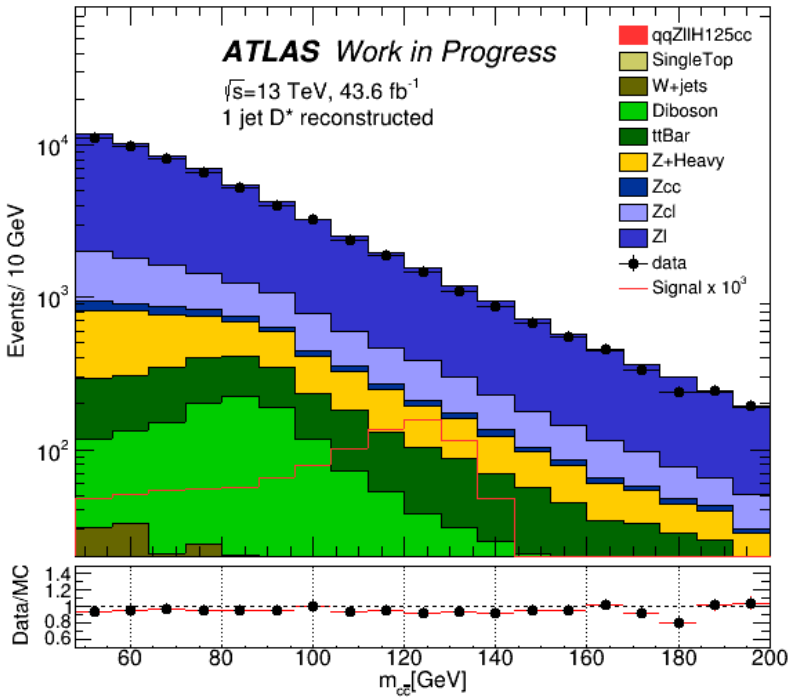


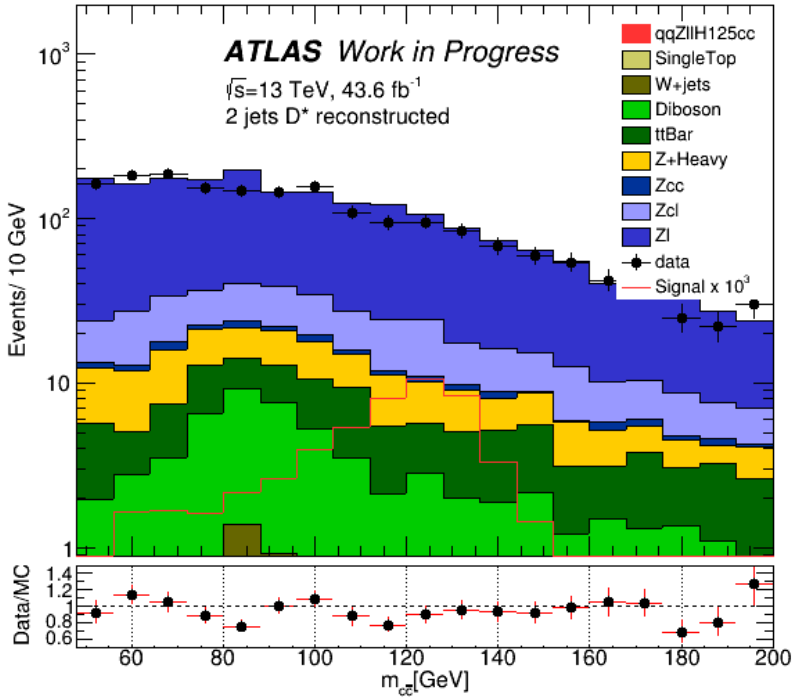
- Major contribution: Z+ light jets
- Signal peaks at 145 *MeV*
- Backgrounds from Zcl and Zcc show a peak too

$\Rightarrow$ Must find a suitable parameter to distinguish the signal from the backgrounds: $m_{c\bar{c}}$

$m_{c\bar{c}}$ Invariant mass

- Jets which pass D^* reconstruction and ΔR cut used to evaluate $m_{c\bar{c}}$
- Analysis split into two:
  1. At least one jet D^* reconstructed
  2. Exactly two jets D^* reconstructed





- Visible signal and background distinction

Moving on

Limit evaluation

- Binned likelihood approach
- Using 15 uniform width bins in the range of $50 GeV < m_{c\bar{c}} < 200 GeV$
- Upper limits calculated on the parameter μ, the ratio of measured signal yield to the prediction from the SM

Case	Expected limit
At least one jet D* tagged	234
Exactly two jets D* tagged	897

- Observed (Expected) Limits from the BDT c-tagging ≤ 150

Yields

Sample	Yield ± (stat.) (Fraction of Total Sim. %)		
	Pre- selection	1 jet D* tagged	2 jets D* tagged
Z + jets	1950258±5035 (95.647%)	256342±1385 (94.473%)	8996 ± 150 (92.323%)
Z + ll	1602670±3581 (78.600%)	211023 ± 989 (77.771%)	7246 ± 112 (74.363%)
Z + cl	190310 ± 734 (9.333%)	25135 ± 211 (9.263%)	999 ± 21 (10.252%)
Z + bl	107915 ± 318 (5.292%)	14464 ± 85 (5.330%)	560 ± 8 (5.747%)
Z + cc	23830 ± 230 (1.168%)	2742 ± 58 (1.010%)	99 ± 6 (1.016%)
Z + bb	18443 ± 103 (0.904%)	1799 ± 24 (0.663%)	52 ± 2 (0.533%)
Z + bc	7090 ± 69 (0.347%)	976 ± 18 (0.359%)	38 ± 1 (0.389%)
$t\bar{t}$	57103 ± 88 (2.818%)	9615 ± 36 (3.543%)	475 ± 8 (4.874%)
Diboson	24729 ± 43 (1.220%)	4264 ± 17 (1.571%)	224 ± 2 (2.301%)
W + jets	2883 ± 229 (0.142%)	537 ± 63 (0.198%)	24 ± 3 (0.247%)
Single top	2349 ± 29 (0.116%)	358 ± 11 (0.132%)	17 ± 2 (0.179%)
$qq \to Z(ll)H(cc)$	14.65 ± 0.02 (0.000%)	3.11 ± 0.01 (0.001%)	0.198±0.002 (0.002%)
Total Sim.	2039004 ± 5431	271338 ± 1514	9744 ± 165
Data	2025900 ± 1423	264229 ± 514	8953 ± 94

Event yields for data and simulation for D^* reconstructed jets - High acceptance rate

Yields

Sample	Yield ± (stat.) (Fraction of Total Sim. %)	
	Pre- selection	2 jets BDT c-tagged
Z + jets	1950258 ± 5035 (95.64%)	472 ±48 (85.81%)
Z + ll	1602670 ± 3581 (78.60%)	23 ±6 (4.18%)
Z + cl	190310 ± 734 (9.33%)	70 ± 8 (12.72%)
Z + bl	107915 ± 318 (5.29%)	9 ± 2 (1.63%)
Z + cc	23830 ± 230 (1.16%)	306 ± 25 (55.63%)
Z + bb	18443 ± 103 (0.90%)	43 ± 4 (7.81%)
Z + bc	7090 ± 69 (0.34%)	19 ± 3 (3.45%)
$t\bar{t}$	57103 ± 88 (2.81%)	51 ± 36 (9.27%)
Diboson	24729 ± 43 (1.22%)	22 ± 1 (0.04%)
W + jets	2883 ± 229 (0.14%)	0.3 ± 0.2 (0.00%)
Single top	2349 ± 29 (0.11%)	1.4 ± 0.4 (0.25%)
$qq \rightarrow Z(ll)H(cc)$	14.65 ± 0.02 (0.00%)	0.162 ± 0.001 (0.02%)
Total Sim.	2039004 ± 5431	550 ± 51
Data	2025900 ± 1423	784 ± 28

Event yields for data and simulation for BDT c-tagged jets -
Low acceptance rate, high signal efficiency with respect to backgrounds

Conclusion

- High value of upper limit may arise due to:
  1. Insufficient efficiency of D^* reconstruction: Limit for 'exactly 2 D^* reconstructed jets' higher than the limit for 'at least one D^* reconstructed jet' case
  2. Loss of events due to fragmentation of charm: D^* reconstruction is performed only on 24% of the charm flavoured jets.

     Visible from double charm background efficiency: Zcc efficiency reduces; Zcl and Zbc efficiency increases after D^* tag $\Leftrightarrow$ effect multiplied for double c-events.

Moving on

Summary and outlook

- Analysis in broadly two parts:
  - Reconstruction of D^* mesons: $m_{K\pi\pi} - m_{K\pi}$ peak at 145.43 ± 0.06 MeV close to the world average of 145.42 ± 0.01 MeV
  - Implementing c-tagging on $H \to c\bar{c}$ data: Obtained an upper limit on signal strength, which is considerably higher than the BDT c-tagging method due to anticipated reasons
- $H \to c\bar{c}$ decays remain one of the challenging decays.
  Current c-tagging algorithms: 41% efficient
- With this work, one can rule out the possibility of using this method for charm tagging for cross-section limit calculation purposes; room still open for the use of D^* reconstruction for data validation techniques

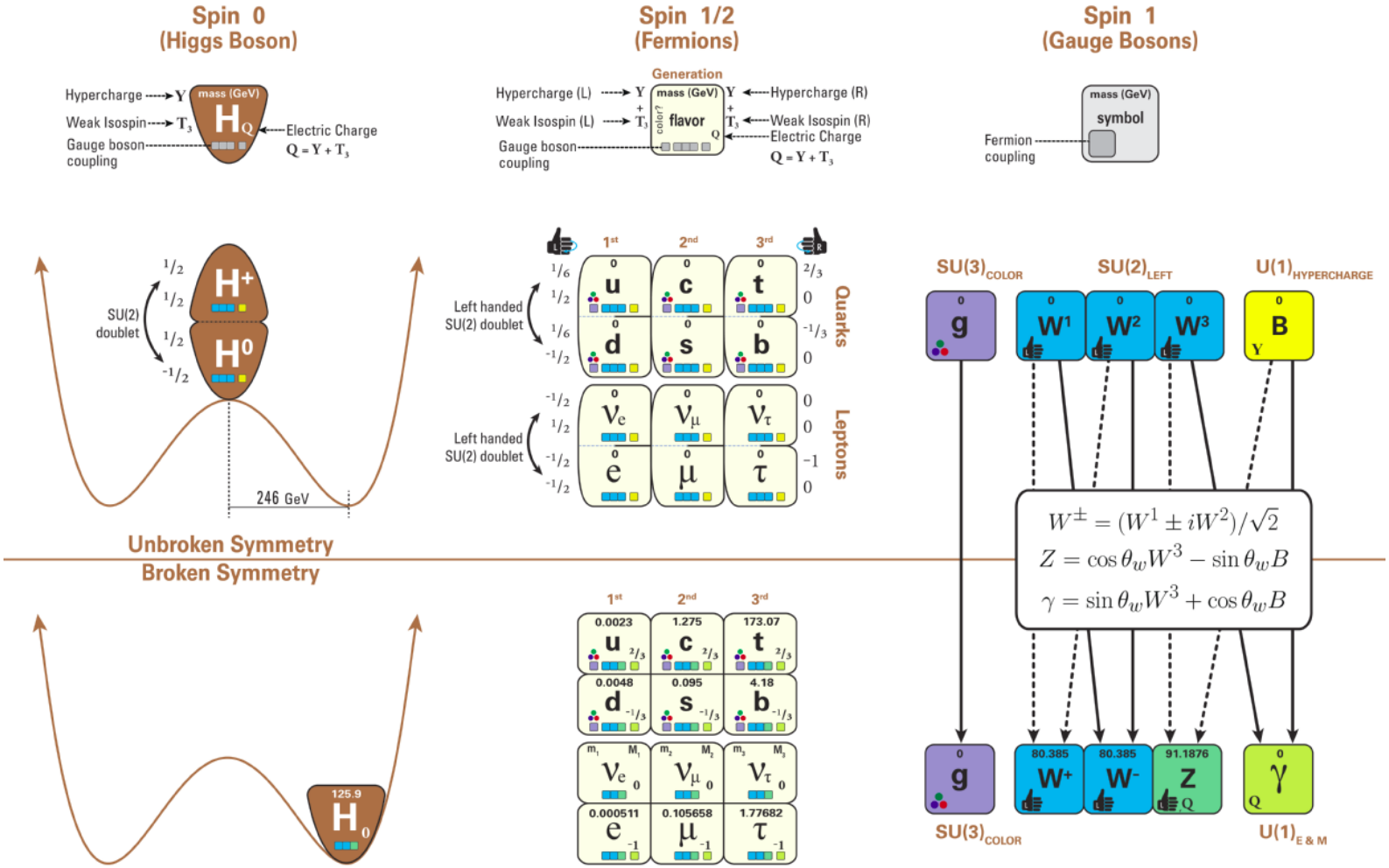
Figure sources

1. Searches for Higgs Boson at CDF; https://www.phys.sinica.edu.tw/~smwang/as_cdf_atlas/cdf/physics/sm_higgs/sm_higgs.html
2. Fermilab Today; https://www.fnal.gov/pub/today/archive/archive_2013/images/Picture_physics_84.jpg
3. Multi-leptons and Top-jets in the Hunt for Gluinos in R-parity Violating Supersymmetry; S. Biswas et al; https://link.springer.com/content/pdf/10.1007%2FJHEP06%282014%29012.pdf
4. Secondary vertex displaced from the primary vertex- diagram by Nazar Bartosik; https://commons.wikimedia.org/wiki/File:B-tagging_diagram.png
5. Tertiary vertex from D-hadron decay; https://indico.cern.ch/event/93145/attachments/1101405/1571210/Sem3_btag.pdf
6. Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run; https://cds.cern.ch/record/2160731/files/ATL-PHYS-PUB-2016-012.pdf
7. Search for the Decay of the Higgs Boson to Charm Quarks with the ATLAS Experiment; Phys. Rev. Lett. 120 (2018) 211802; https://arxiv.org/pdf/1802.04329.pdf

References

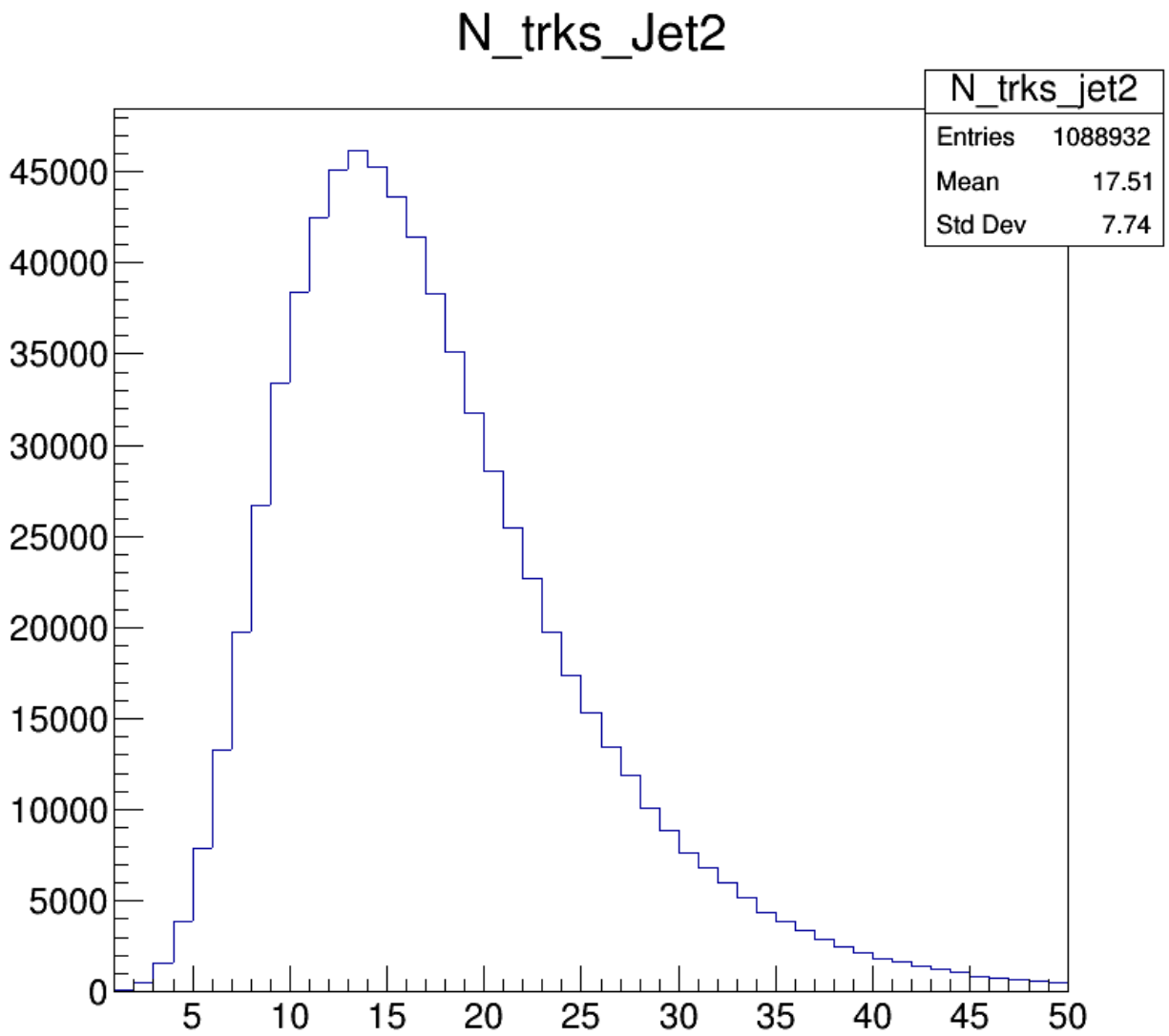
- The ATLAS Collaboration, *Search for the decay of the Higgs boson to charm quarks with the ATLAS experiment*; DOI:10.1103/PhysRevLett.120.211802, 2018.
- ATLAS flavor tagging group; *b- tagging in ATLAS* `https://indico.cern.ch/event/242419/contributions/520667/attachments/412165/572722/B-tag2012.pdf`
- The ATLAS Collaboration, *Search for $H \to c\bar{c}$ decays in associated ZH production*; ATLAS Note HIGG-2017-01
- The ATLAS Collaboration, *b-jet tagging calibration on c-jets containing D^{*+} mesons*; ATLAS-CONF-2012-039
- CMS Collaboration, *Measurement of associated production of a W boson and a charm quark in proton-proton collisions at $\sqrt{s} = 13\,TeV$*; Eur. Phy. J. C(2019) 79:269

Higgs particle

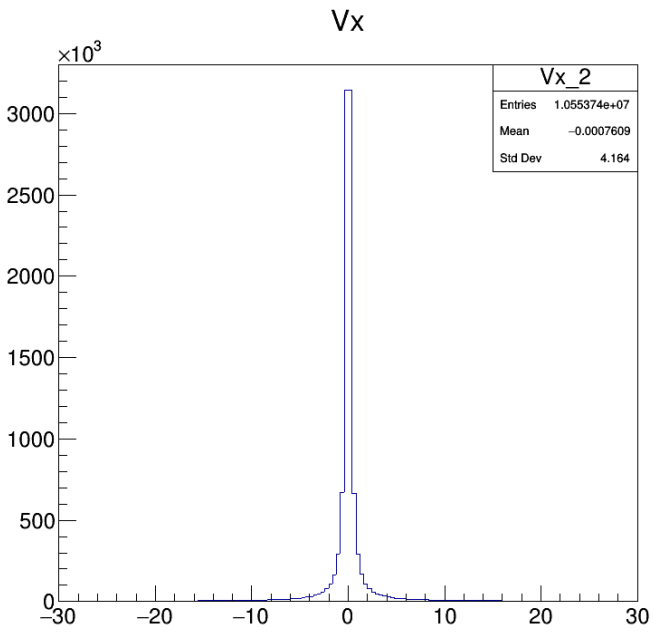


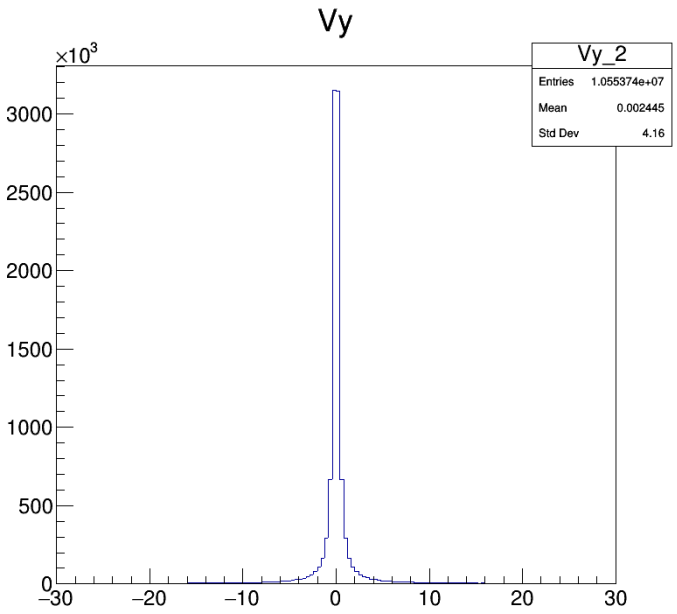
By Latham Boyle, retrieved from:
https://commons.wikimedia.org/wiki/File:Standard_Model_Of_Particle_Physics,_Most_Complete_Diagram.jpg

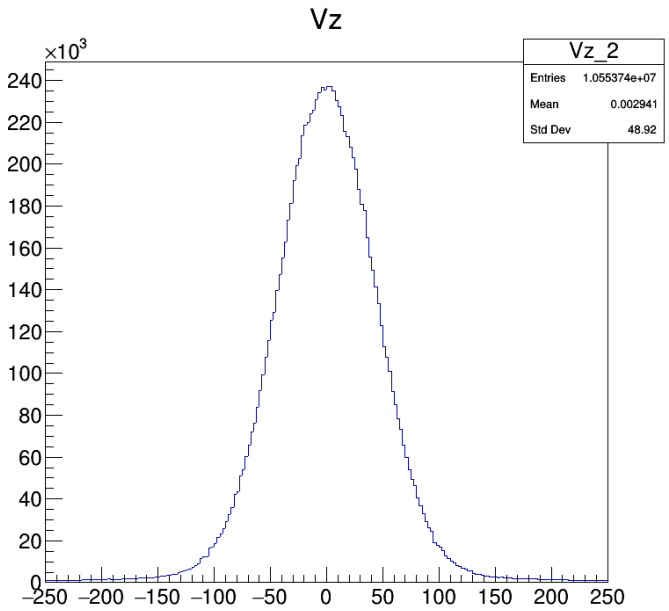
Tracks in jet

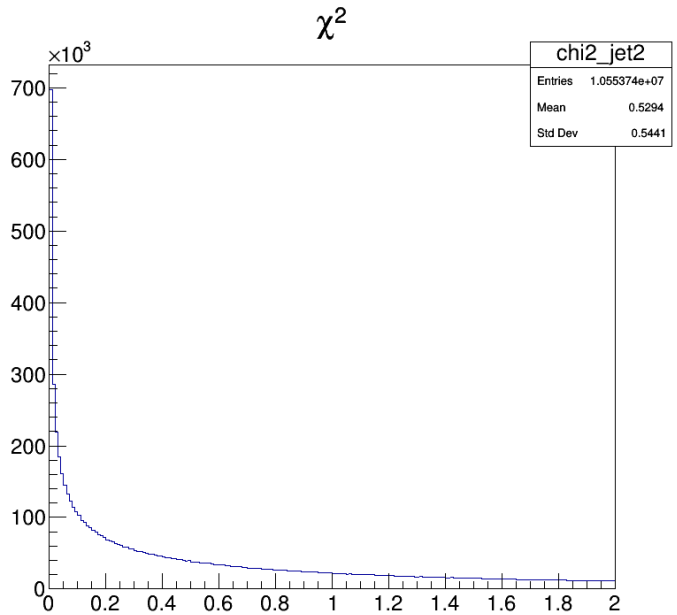


Vertex obtained









Distance of π_{slow} from D^0 vertex