Introduction to b-Tagging

HEP Postgraduate Lecture Course

21.11.17

Andrew Bell

Motivation

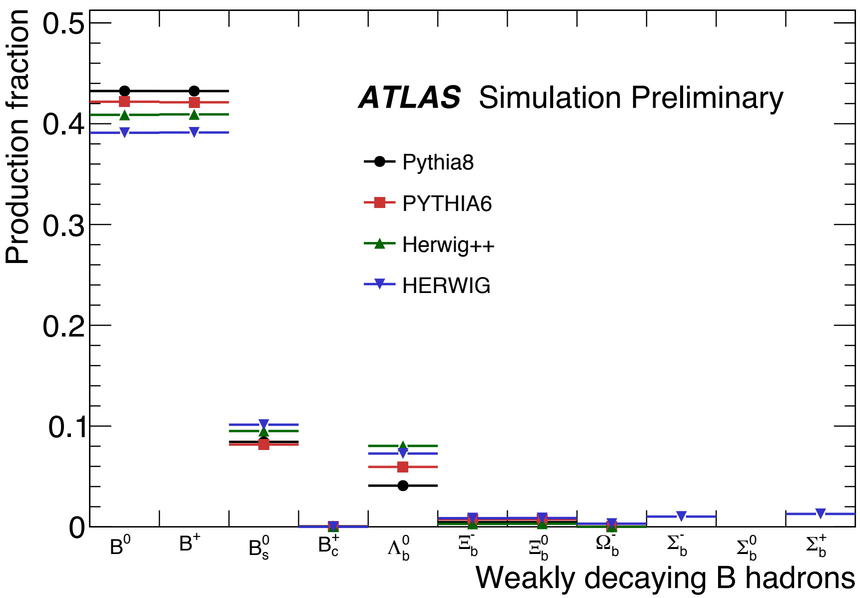
- Many interesting physics signatures produce a final-state b-quark:
  - Top decays ($|V_{tb}| >> |V_{ts}|$), Higgs→bb, BSM physics searches, precision standard model measurements etc.

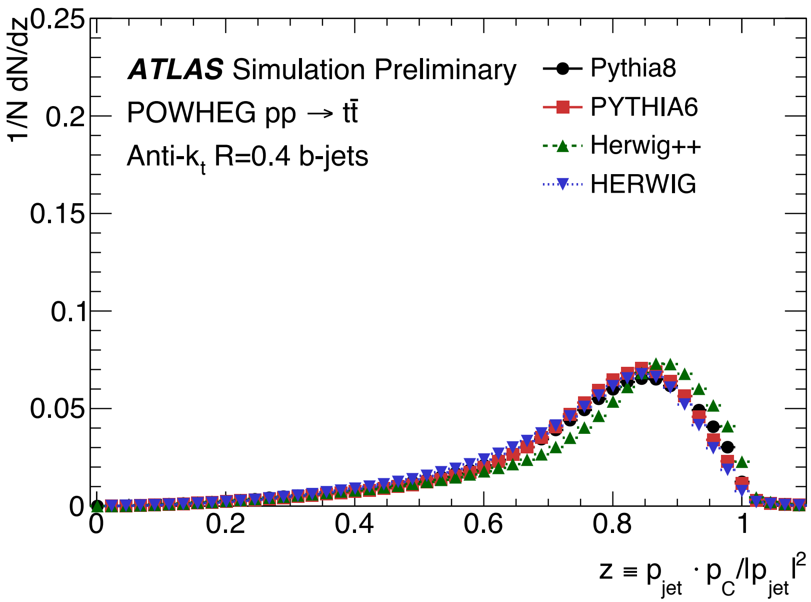
- Final state b-quark will hadronise and shower
  - Reconstructed as a jet
  - What properties can we use to differentiate these jets from others?

- Truth flavour label assigned in simulation using ΔR matching between hadron and jet:
  - If b-hadron present, label as "b-jet"
  - Else if no b-hadron but c-hadron present, label as "c-jet" (c-hadrons have similar properties to b-hadrons)
  - Else label as a "light-flavour jet"

b-Hadron Properties

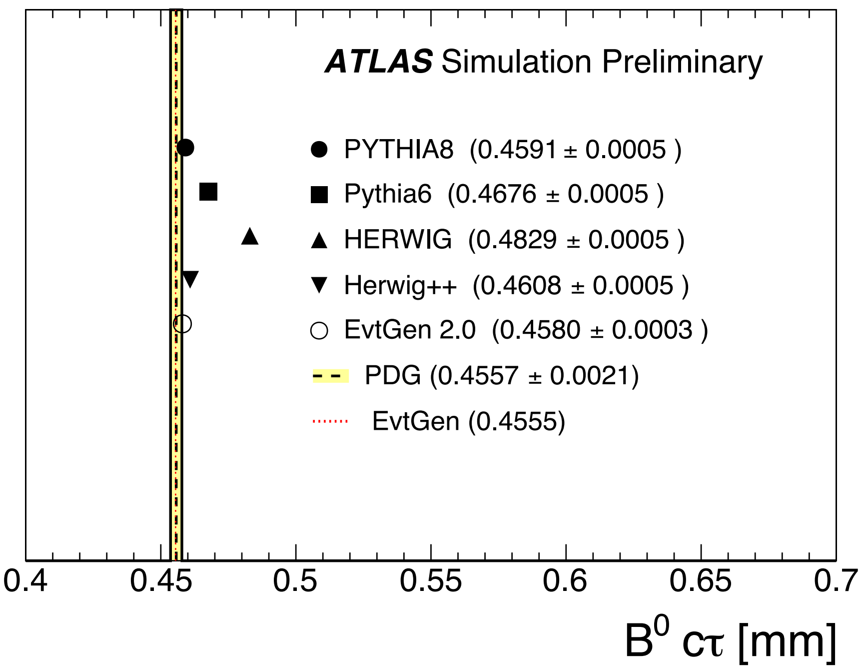
- Final state b-quark fragments (production of hadrons from quarks) to excited b-hadrons:
  - B*/B** ~87% of the time
- These decay strongly or electromagnetically into a stable b-hadron (with a few additional particles) (left figure)
- b-quark fragmentation is quite hard:
  - ~70% of b-quark energy goes to b-hadron (right figure)
- b-hadrons typically decay to produce ~5 charged stable decay products
  - c-hadrons typically decay to produce ~2 charged stable decay products

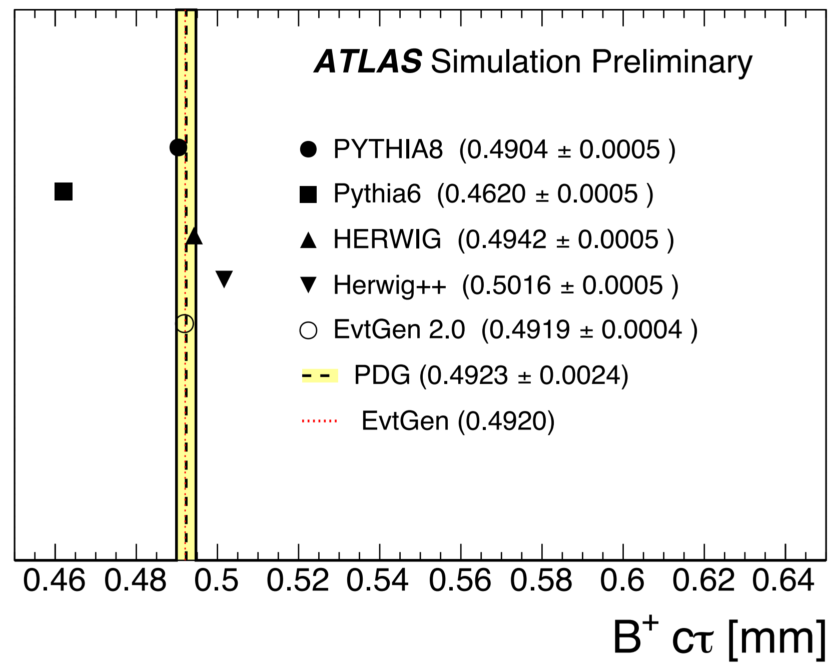




b-Hadron Lifetimes

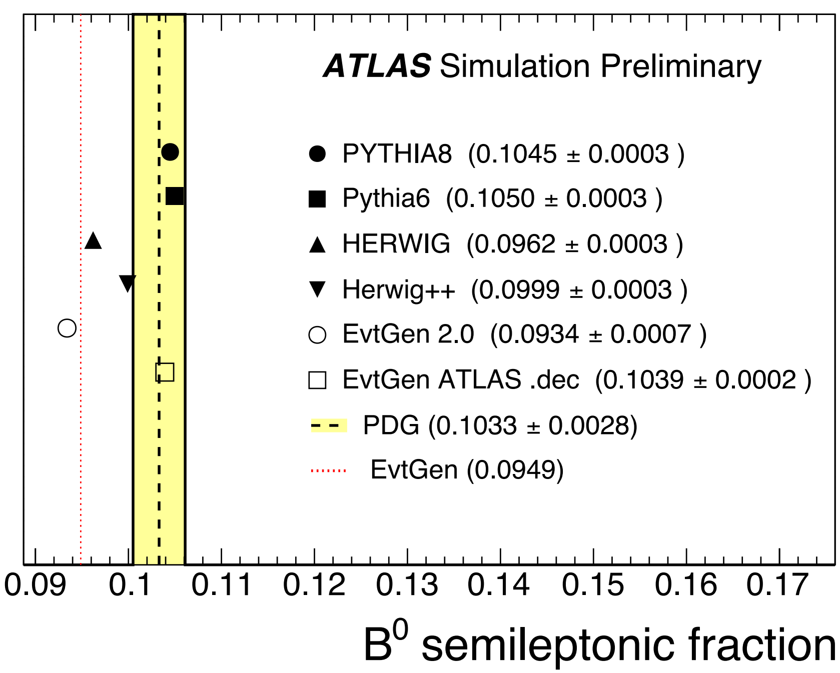
- Lifetime of b-hadron is typically ~1.5 ps
- Plots below compare decay distances of b-hadrons in a number of commonly used MC generators
- EvtGen is used throughout ATLAS as MC "afterburner":
  - Provide updated lifetime and decay information
  - More realistic modelling and improve the description of the particle decay modes and multiplicities
- Equivalent plots for c-hadrons in back-up

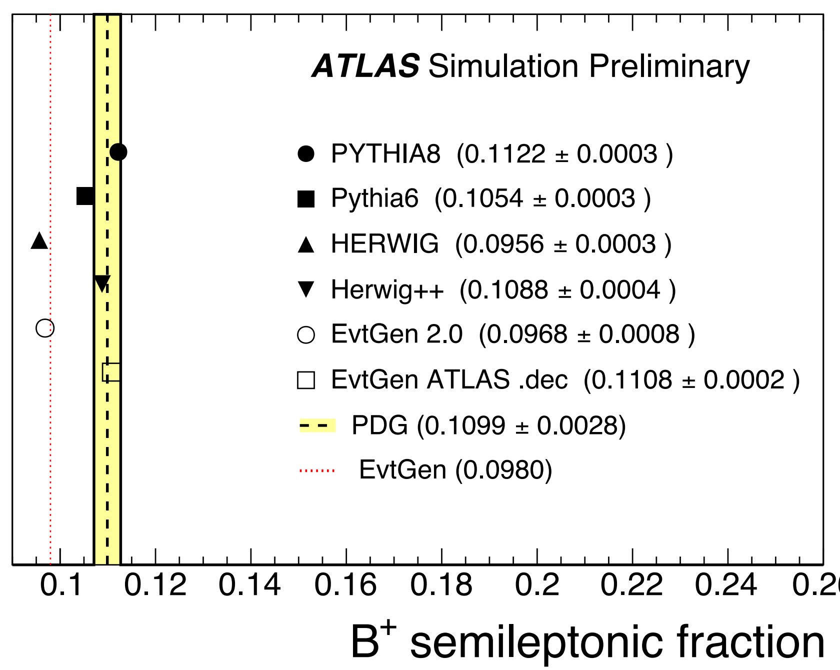




b-Hadron Lifetimes

- b-hadrons decay semileptonically ~10% of the time (i.e. produces an electron or muon)
- Plots below compare semileptonic fractions of b-hadrons in a number of commonly used MC generators
- EvtGen is used throughout ATLAS as MC “afterburner”:
  - Provide updated lifetime and decay information
  - More realistic modelling and improve the description of the particle decay modes and multiplicities
- Equivalent plots for c-hadrons in back-up

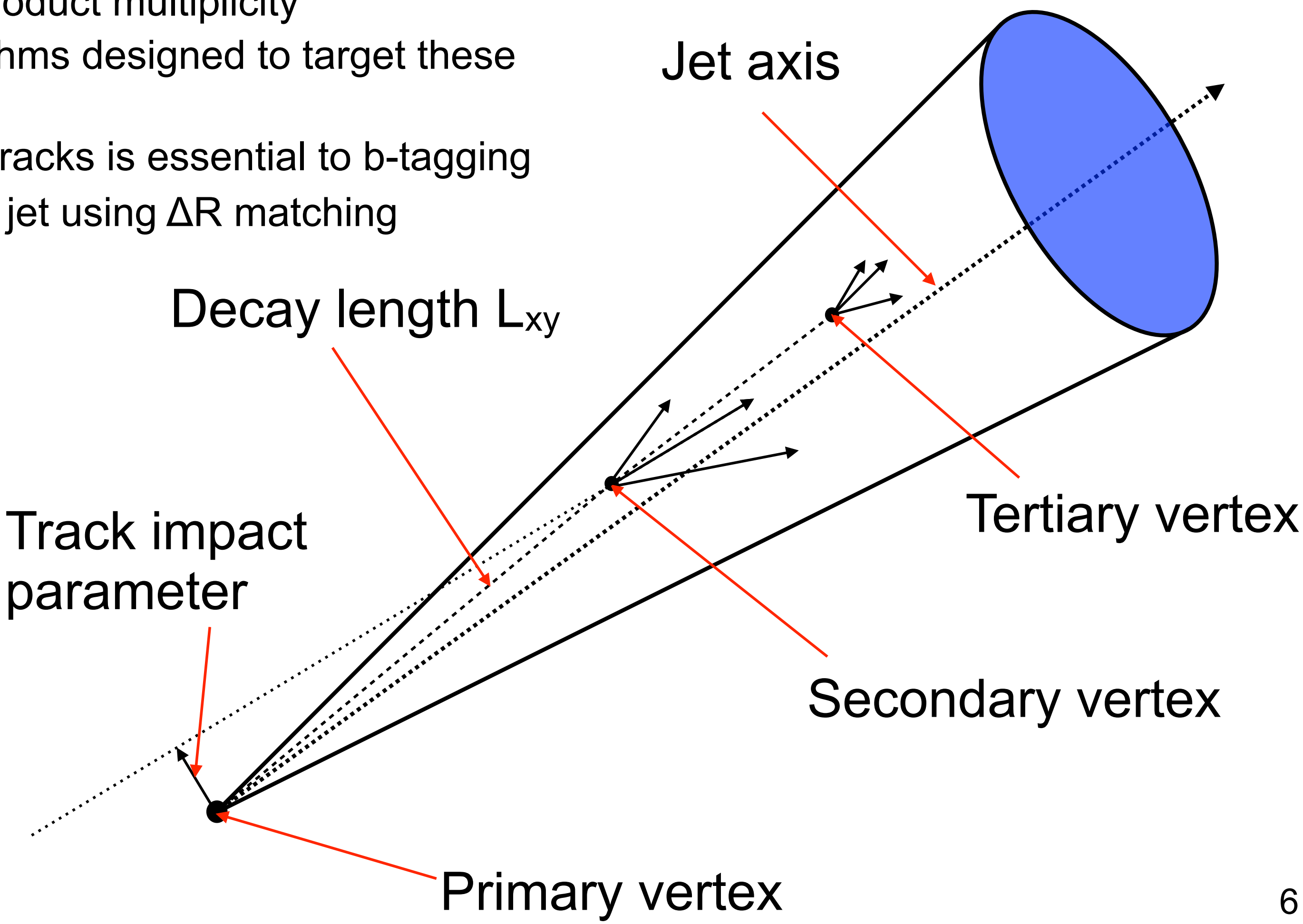




Jet Properties

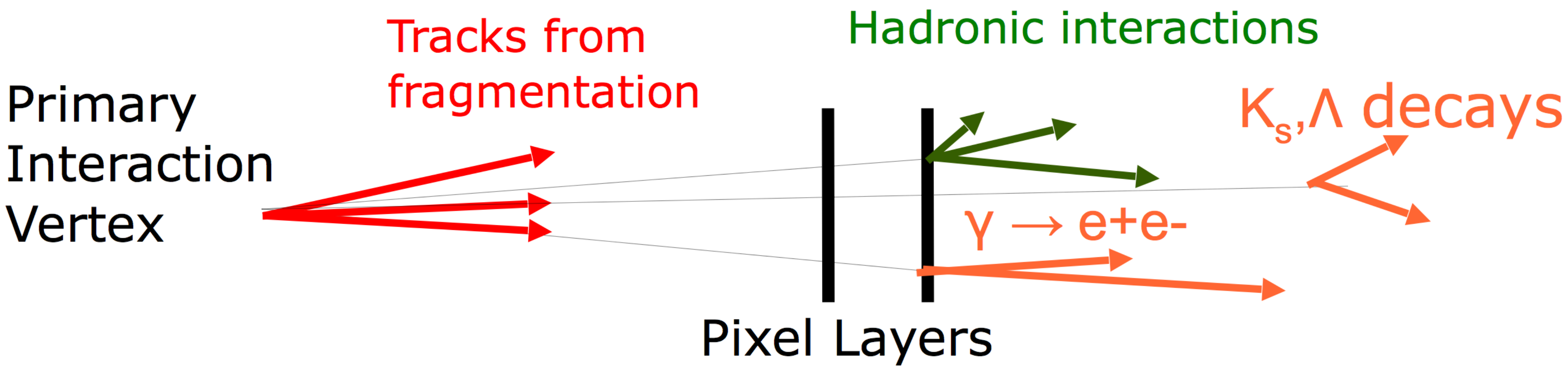
- b-hadrons have a number of useful properties:
  - Longer lifetimes (~1.5 ps)
  - b-hadron decays to a c-hadron ($|V_{cb}| >> |V_{ub}|$)
  - Larger mass (~5 GeV)
  - High decay product multiplicity
- 3 "baseline" algorithms designed to target these properties
- Reconstruction of tracks is essential to b-tagging
  - Associated to jet using ΔR matching

> » For a 30 GeV b-hadron
>
> » $\beta\gamma c\tau$ ~ 5 mm - measurable!



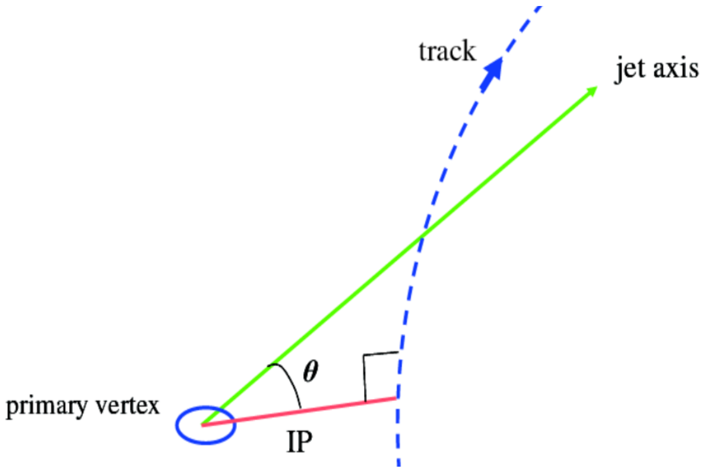
Light-flavour Jet Properties

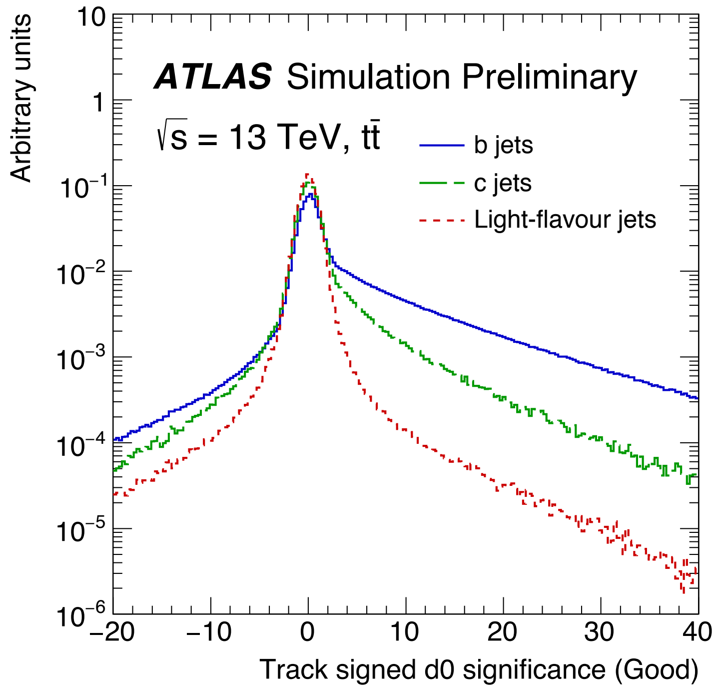
- In a light-flavour jet, most tracks come directly from quark fragmentation
- Can sometimes result in a displaced vertex and look like a b-jet:
  - Interactions with the detector material
  - Photon conversions
  - Long-lived particles (K_s/Λ)
  - Badly measured tracks

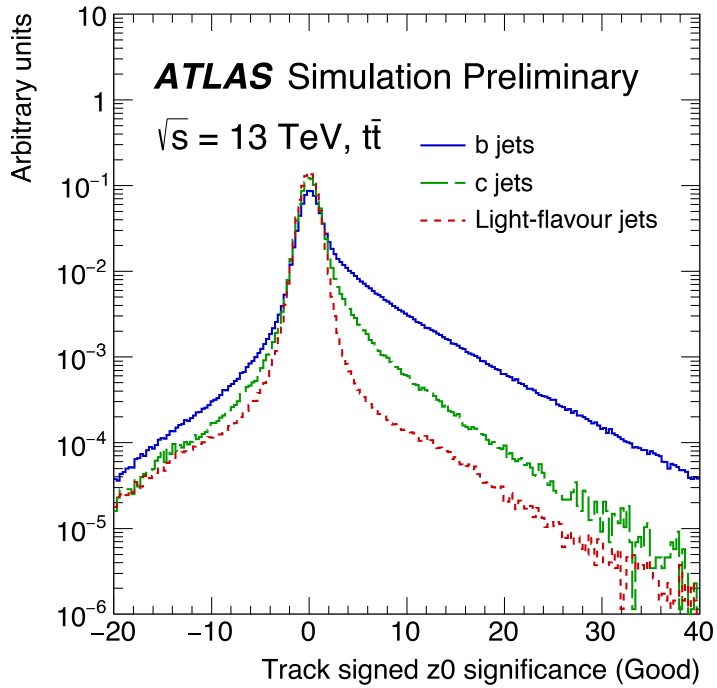


Impact Parameter Algorithms

» Impact parameters (IP) defined as signed point of closest approach in longitudinal and transverse plane
  » IP Significance is the IP divided by its uncertainty
  » IP is signed: positive if crosses jet axis in front of primary vertex (otherwise negative)
» Construct reference templates (PDFs) for b-, c- and light-flavoured jets
» Log likelihood ratio of probability for each jet flavour gives best separation (Neyman-Pearson lemma)
  » Assume each track is uncorrelated
» IP2D uses just transverse (d_0) IP (less susceptible to pileup)
» IP3D uses transverse (d_0) and longitudinal (z_0) IP



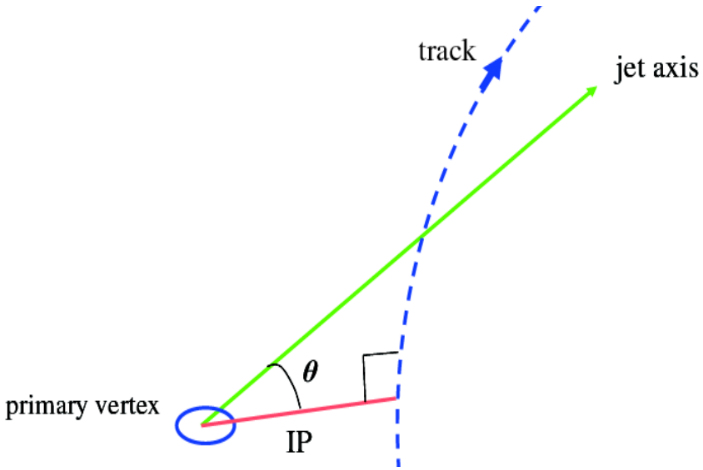




Impact Parameter Algorithms

UCL

- Impact parameters (IP) defined as signed point of closest approach in longitudinal and transverse plane
  - IP Significance is the IP divided by its uncertainty
  - IP is signed: positive if crosses jet axis in front of primary vertex (otherwise negative)
- Construct reference templates (PDFs) for b-, c- and light-flavoured jets
- Log likelihood ratio of probability for each jet flavour gives best separation (Neyman-Pearson lemma)
  - Assume each track is uncorrelated
- IP2D uses just transverse (d_0) IP (less susceptible to pileup)
- IP3D uses transverse (d_0) and longitudinal (z_0) IP



Product over all associated tracks

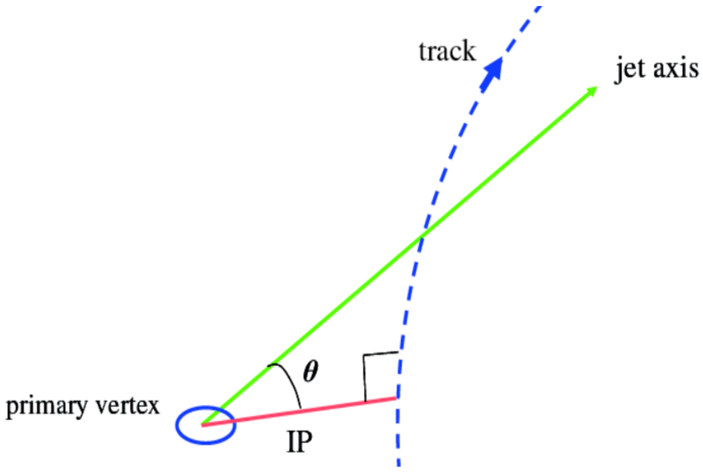
$$\log(P_i/P_j) = \log\left(\frac{\prod_{m=1}^{N} PDF_i(IP_m)}{\prod_{m=1}^{N} PDF_j(IP_m)}\right)$$

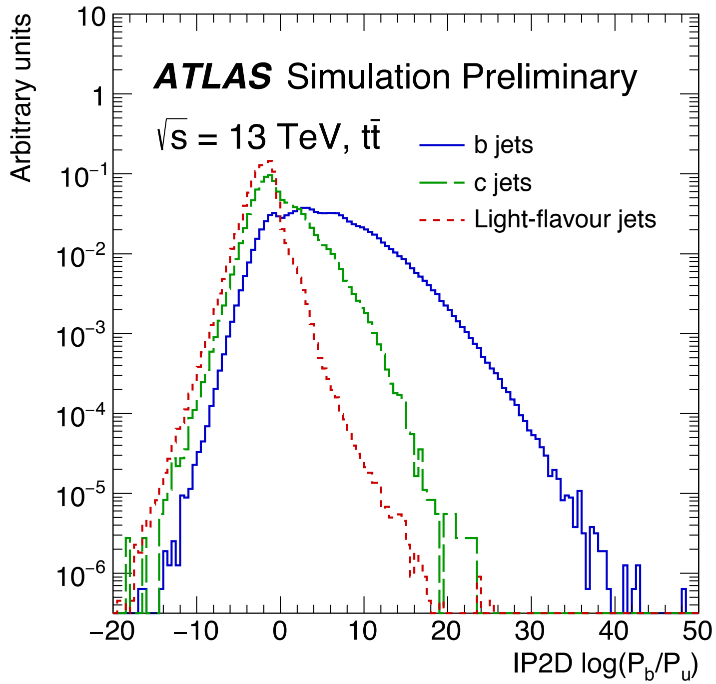
Probability for jet to have flavour j

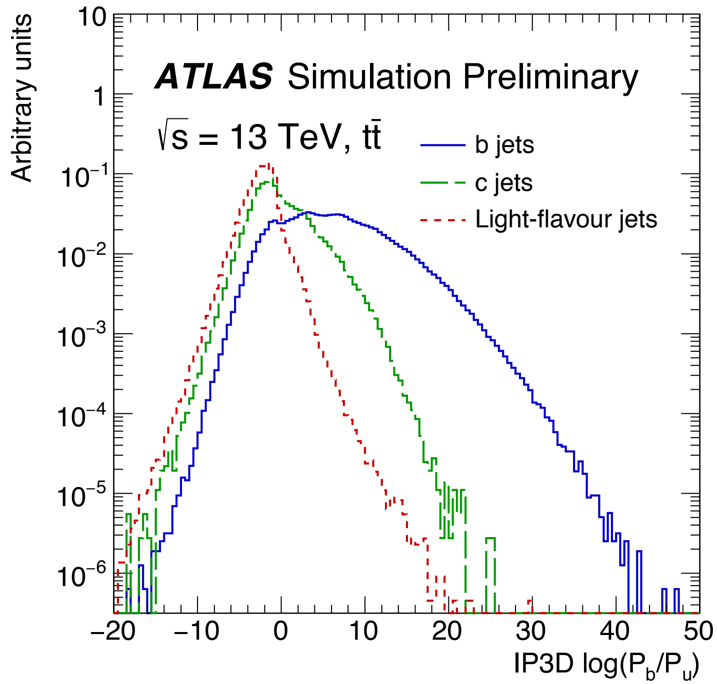
PDF for track of flavour j, with IP_m

Impact Parameter Algorithms

- Impact parameters (IP) defined as signed point of closest approach in longitudinal and transverse plane
  - Significance is the IP divided by its uncertainty
  - IP is signed: positive if crosses jet axis in front of primary vertex (otherwise negative)
- Construct reference templates (PDFs) for b-, c- and light-flavoured jets
- Log likelihood ratio of probability for each jet flavour gives best separation (Neyman-Pearson lemma)
  - Assume each track is uncorrelated
- IP2D uses just transverse (d_0) IP (less susceptible to pileup)
- IP3D uses transverse (d_0) and longitudinal (z_0) IP

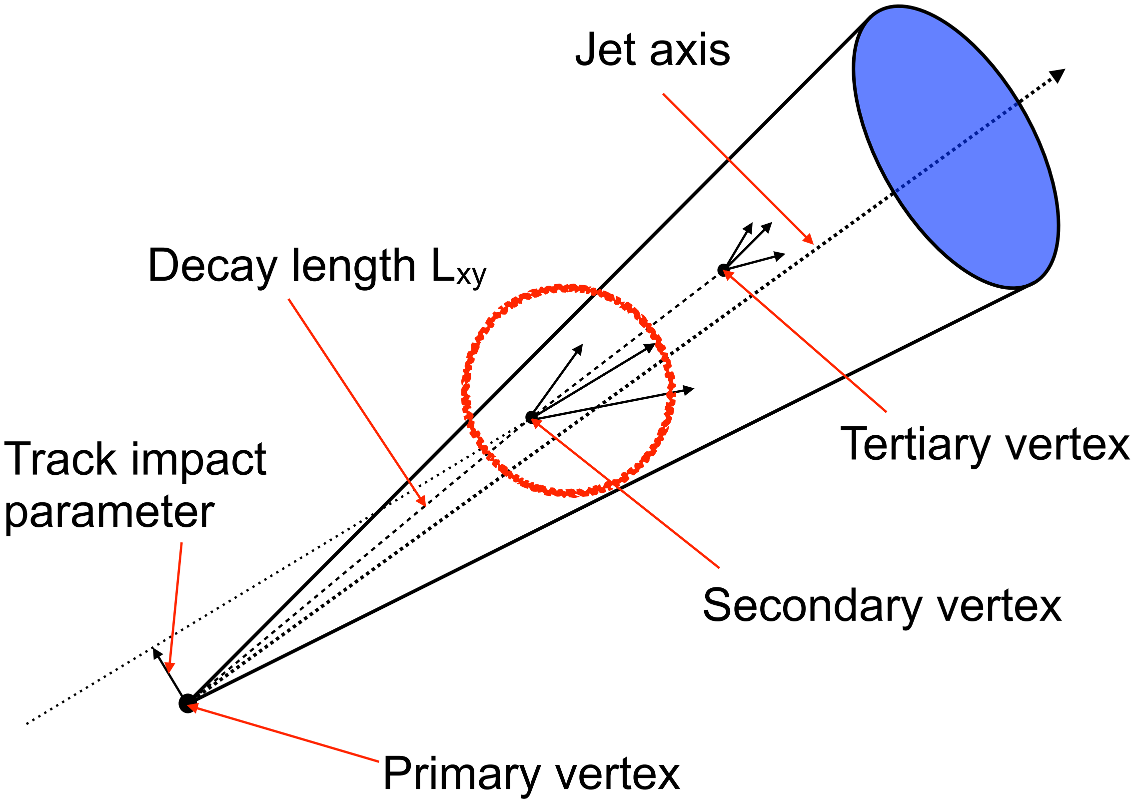






Secondary Vertex (SV) Algorithms

- » Secondary vertex algorithms aim to reconstruct the secondary decay point using tracks
- » Firstly, reconstruct all 2-track vertices
- » Combine into one secondary vertex (removing tracks which are not consistent with the SV)
- » Many useful properties:
  - » Decay length (significance)
  - » Vertex Mass
  - » Number of tracks etc.



Secondary Vertex (SV) Algorithms

» Secondary vertex algorithms aim to reconstruct the secondary decay point using tracks

» Firstly, reconstruct all 2-track vertices

» Combine into one secondary vertex (removing tracks which are not consistent with the SV)

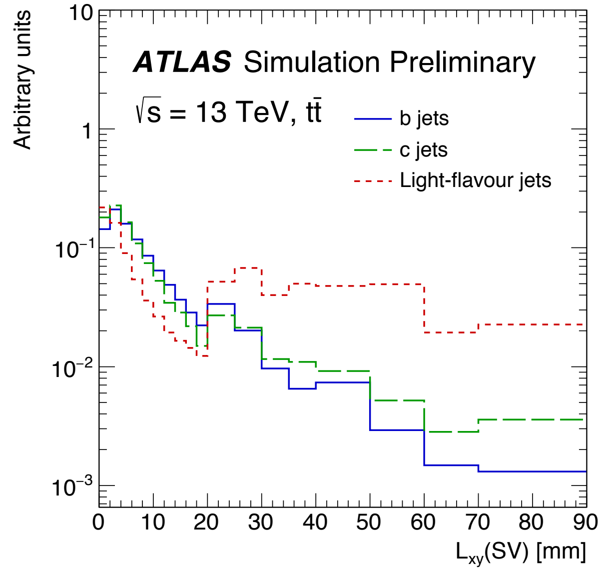
» Many useful properties:

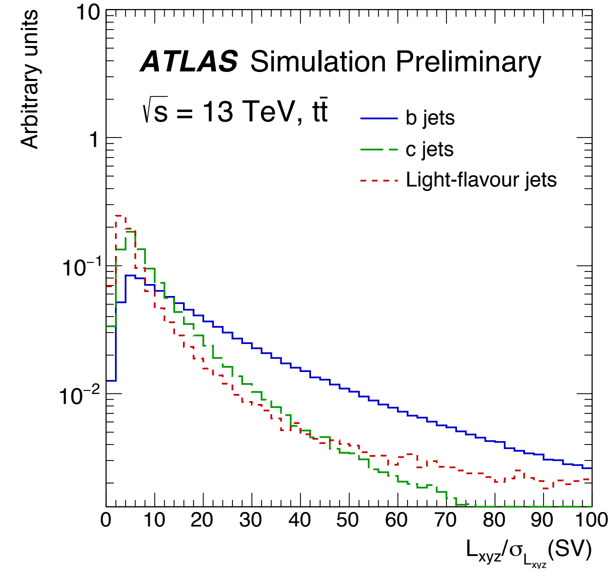
  » Decay length (significance)

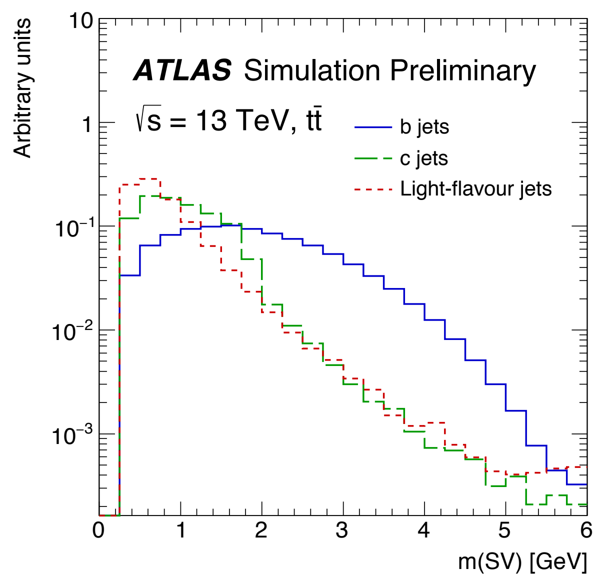
  » Vertex Mass

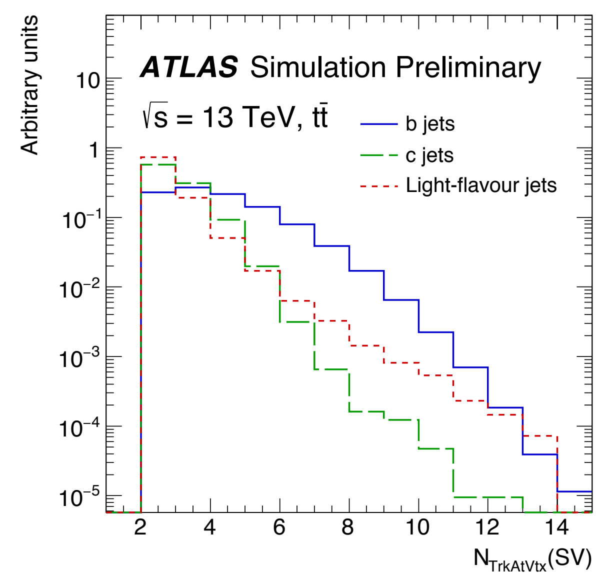
  » Number of tracks

  etc.









Decay Chain Reconstruction Algorithms

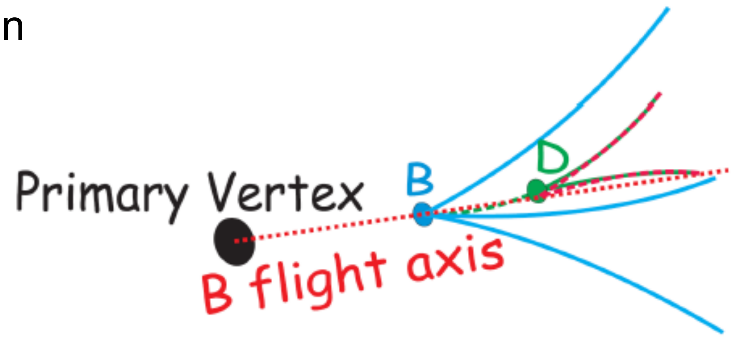
» $|V_{cb}| >> |V_{ub}| \rightarrow$ often produce c-hadron from b-hadron decay

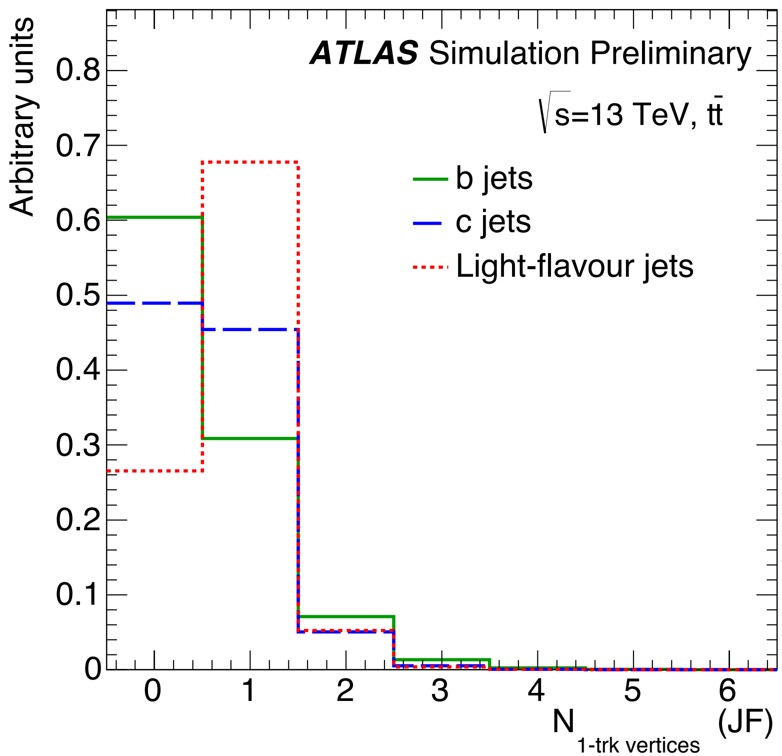
» Additional tertiary vertex

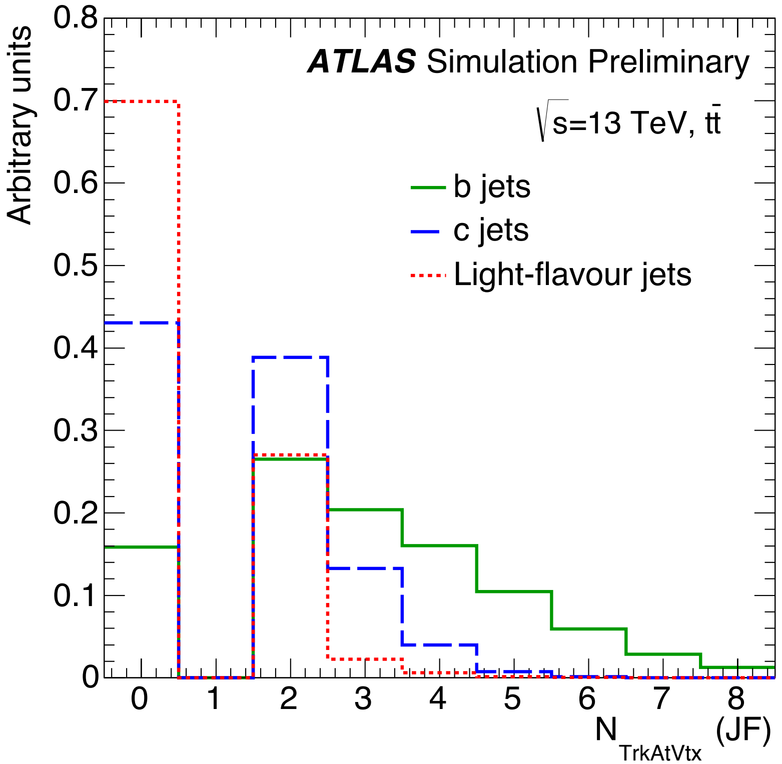
» JetFitter algorithm reconstructs all decay vertices (assuming they lie on the same flight path)

» Can even reconstruct single track vertices

  » Vertex mass, decay length significance, number of vertices with at least 2 tracks etc.







Combining Taggers: Multivariate Algorithms

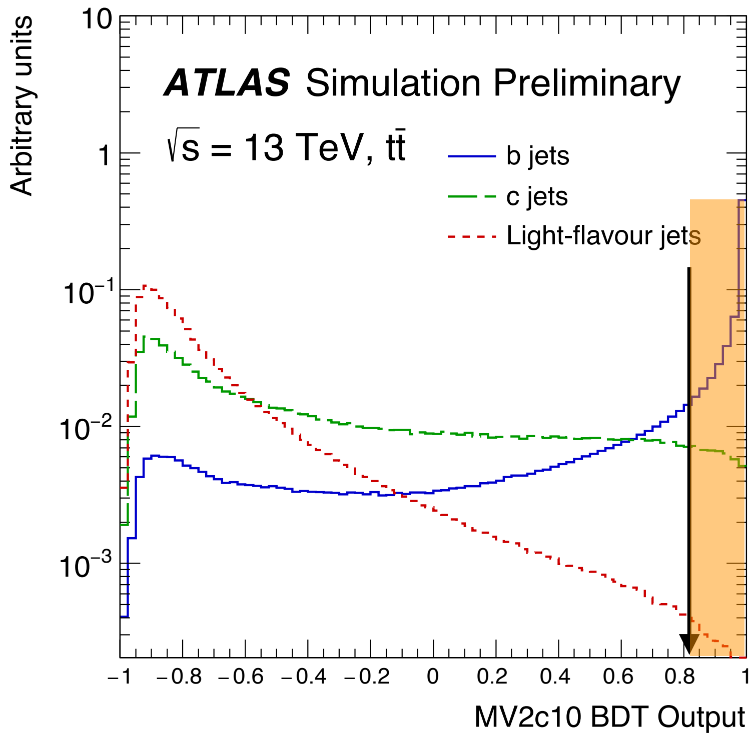
- Three baseline algorithms provide b-jet discrimination → provide a number of weakly correlated variables
- Combine output of three baseline algorithms using a Boosted Decision Tree (BDT)
  - Total of 24 input variables
- Algorithm known as MV2c
- Train using $t\bar{t}$ events, while controlling the c-/light-flavour jet background
  - MV2cXX means trained on a sample with XX% c-jets (100-XX% light-flavour jets)
  - Exposes training to different amounts of each jet type
  - Need to balance light-flavour and c-jet performance
- Efficiency defined as:

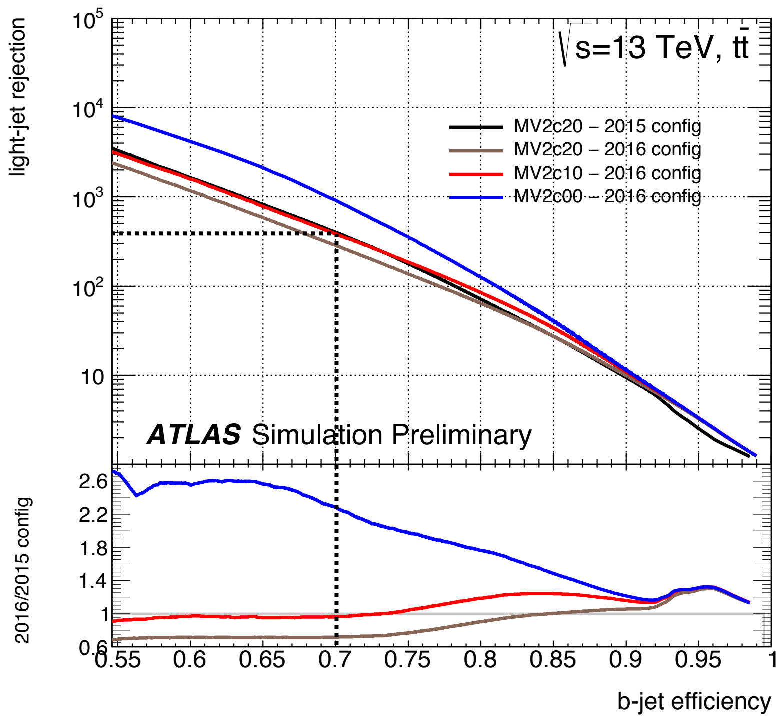
$$\varepsilon_j = \frac{\text{Number of jets of flavour } j \text{ passing cut}}{\text{Number of jets of flavour } j}$$

- Used to evaluate performance of an algorithm at a set "Working Point" (WP)
- Rejection is defined as 1/efficiency

Multivariate Algorithms

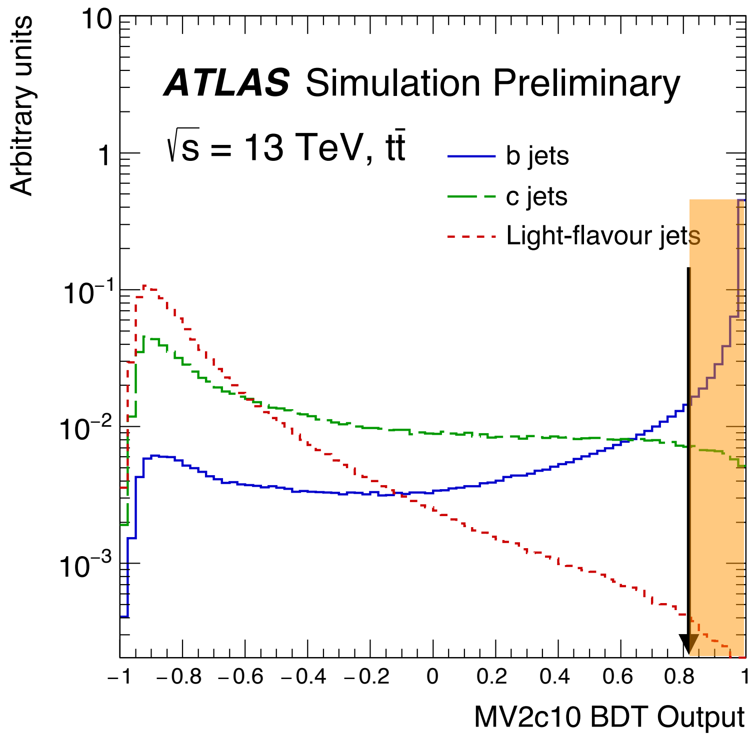
- Light-flavour and c-jet rejection as a function of b-jet tagging efficiency for a number of training configurations
- Larger area under curve → increased performance
- MV2c10 chosen to balance performance of jet flavours
  - Current default throughout ATLAS

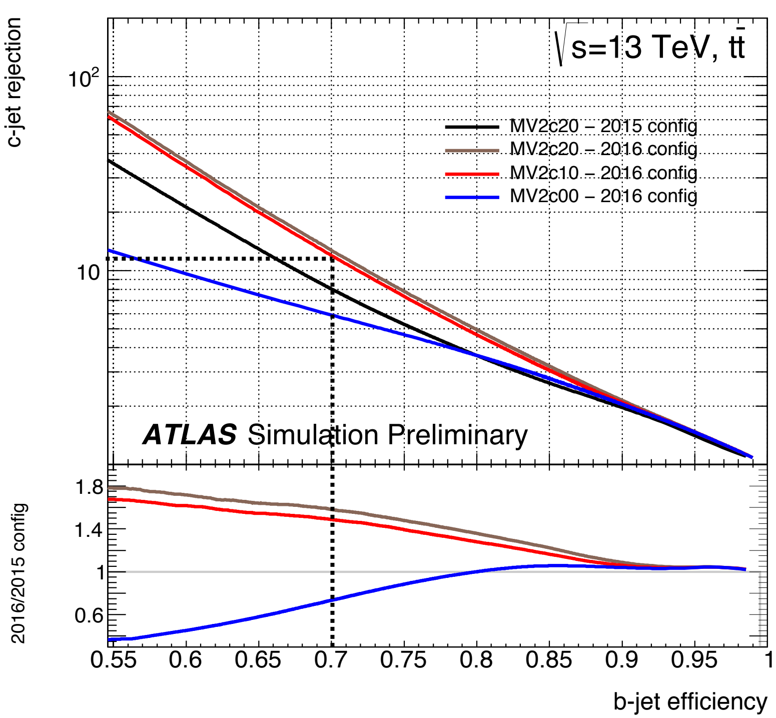




Multivariate Algorithms

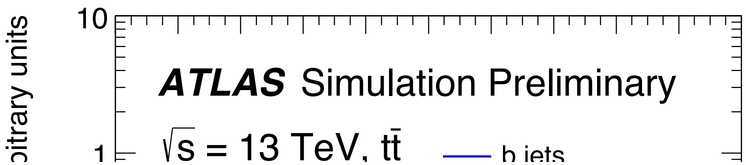
- Light-flavour and c-jet rejection as a function of b-jet tagging efficiency for a number of training configurations
- Larger area under curve → increased performance
- MV2c10 chosen to balance performance of jet flavours
  - Current default throughout ATLAS

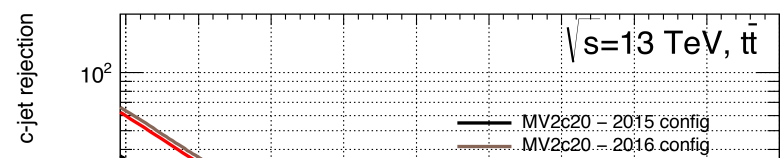




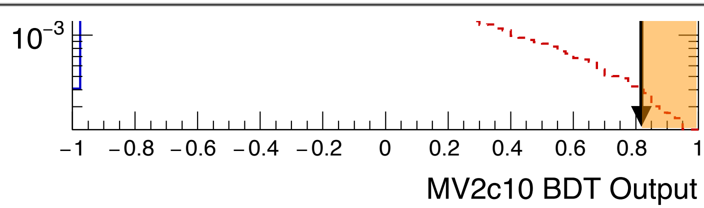
Multivariate Algorithms

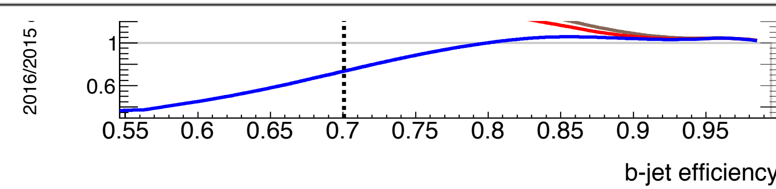
- Light-flavour and c-jet rejection as a function of b-jet tagging efficiency for a number of training configurations
- Larger area under curve → increased performance
- MV2c10 chosen to balance performance of jet flavours
  - Current default throughout ATLAS



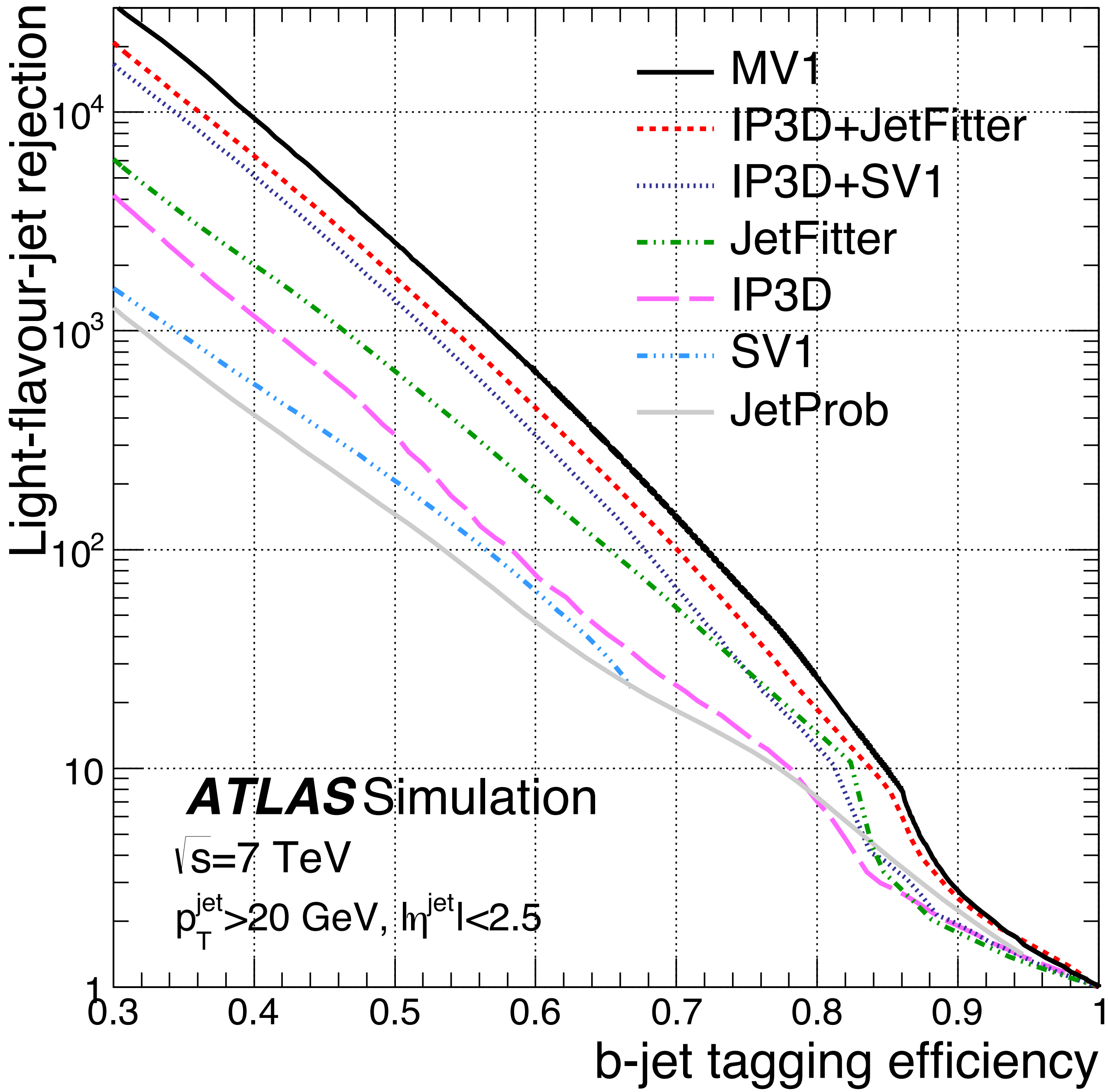


b-jet efficiency [%]	MV2c10 cut	c-jet rejection	τ-jet rejection	light-flavour jet rejection
60.0	0.94	34.5	184.0	1538.8
70.0	0.82	12.2	54.7	381.3
77.0	0.65	6.2	22.0	134.3
85.0	0.18	3.1	8.2	33.5



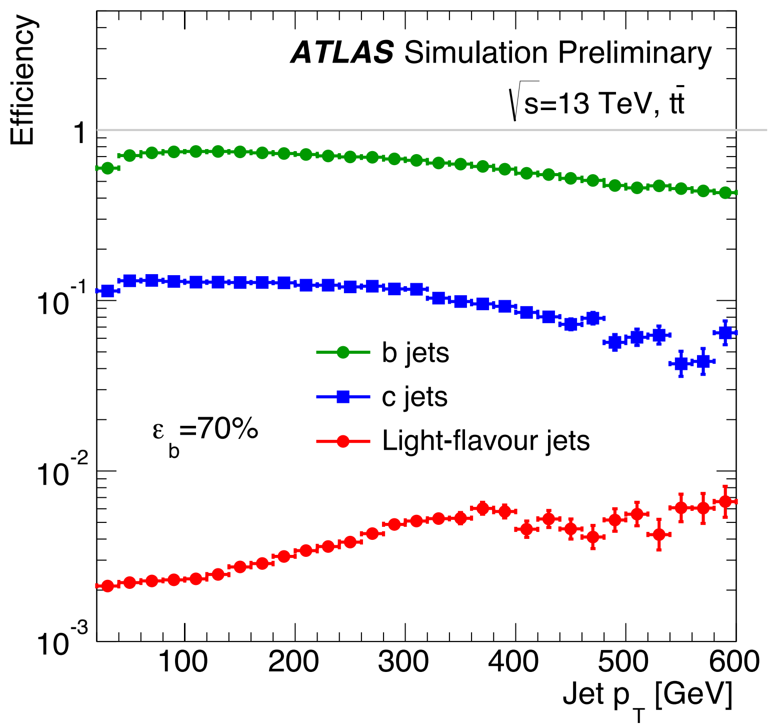


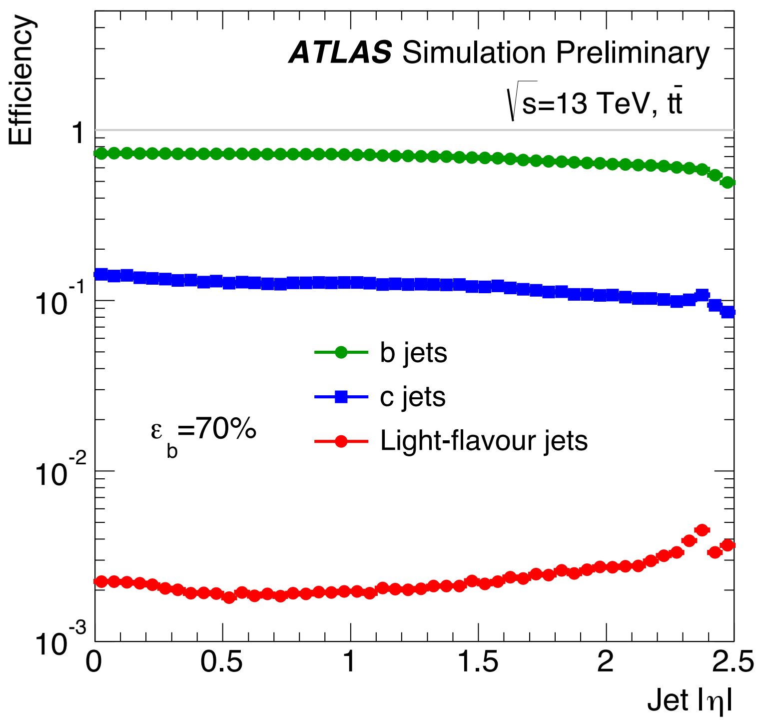
Individual Tagger Performance



Differential Performance

- b-tagging efficiency is dependent on jet and event variables:
  - Performance shown for 70% WP
  - Efficiency drops at low p_T due to reduced IP resolution
  - Drops at high p_T due to b-hadron decaying after first ID layer (reduced IP resolution) and tracks becoming more collimated
- At high |η|, more material interactions within the ID, reducing IP resolution





Calibrations

UCL

- Up to now, all performance studies/training are from simulation
- How well do we know the performance in data?
  - Need to correct for detector and modelling effects
- Apply scale factors to correct performance in MC to data
  - Applied to event weight (for each jet)

Scale factor →

$$\kappa_j = \frac{\varepsilon_j^{\mathrm{data}}}{\varepsilon_j^{\mathrm{MC}}}$$

Efficiency from data

Efficiency from simulation

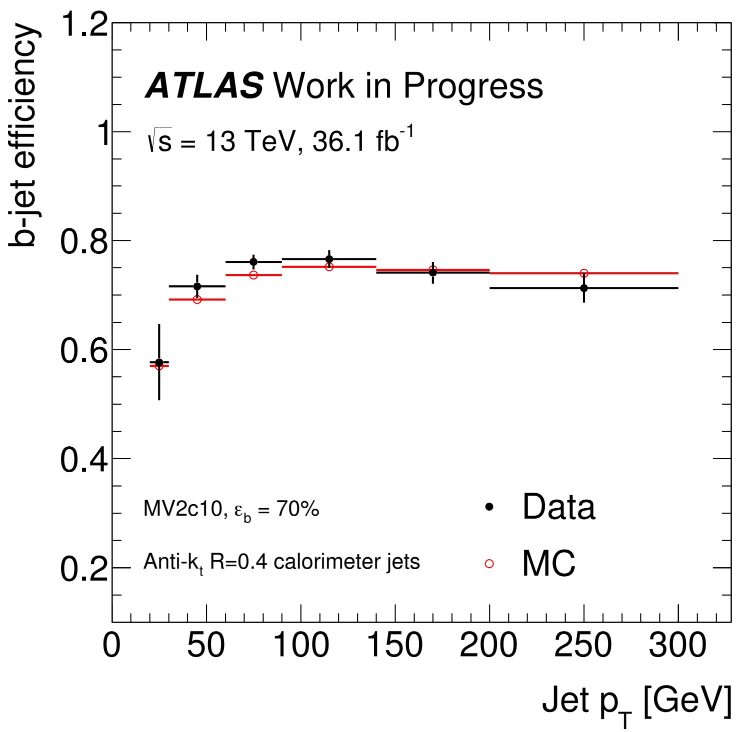
- Select data sample pure in one jet flavour
- Measure efficiency in data for a chosen jet flavour (j)
  - Efficiency in MC is known from truth information
- Derive scale factor as a function of jet p_T and $|\eta|$
- Efficiency measurements conducted at a set working point

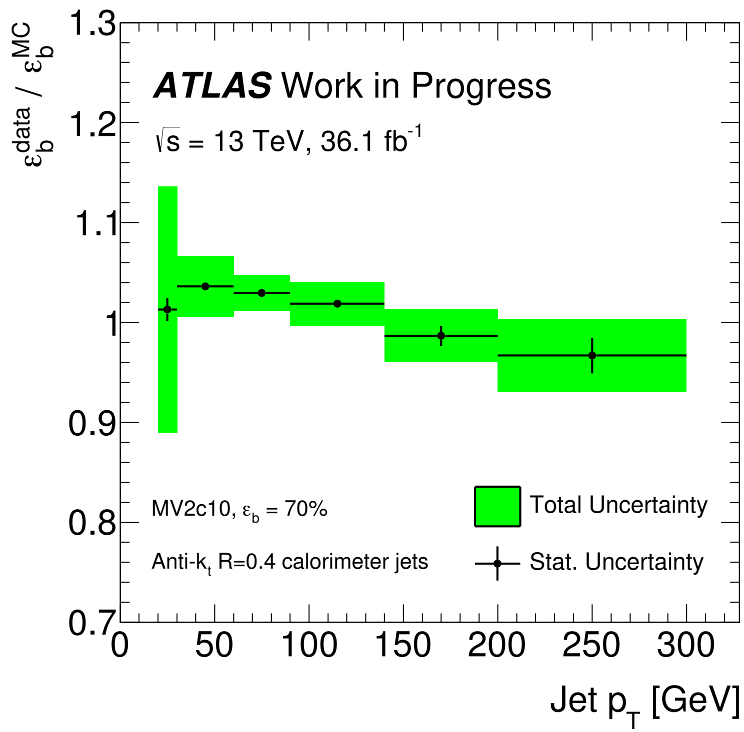
$t\bar{t}$ Calibration

UCL

- b-jets tagging efficiency measured in data using di-leptonic $t\bar{t}$ events
- ~70% pure in b-jets, almost no c-jet contamination
  - Extract b-jet efficiency from fit to data
- Dominated by systematic uncertainties from modelling of $t\bar{t}$ background:
  - Systematics arising from b-tagging are a key factor in a many analyses
  - ~2-5% uncertainties over majority of spectrum
  - Largest experimental systematic uncertainty in VH(→bb) analysis
- Consistent with unity for b-jets, but not always the case!
  - **Light-flavour jets can have SF ~ 2**

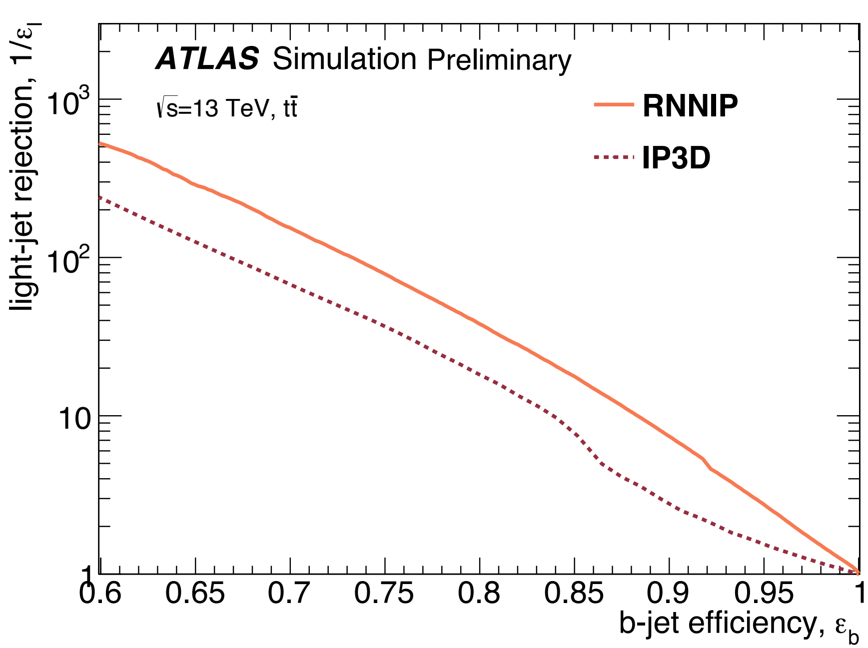
Source of uncertainty		σ_μ
Total		0.39
Statistical		0.24
Systematic		0.31
b-tagging	*b*-jets	0.09
	c-jets	0.04
	light jets	0.04
	extrapolation	0.01

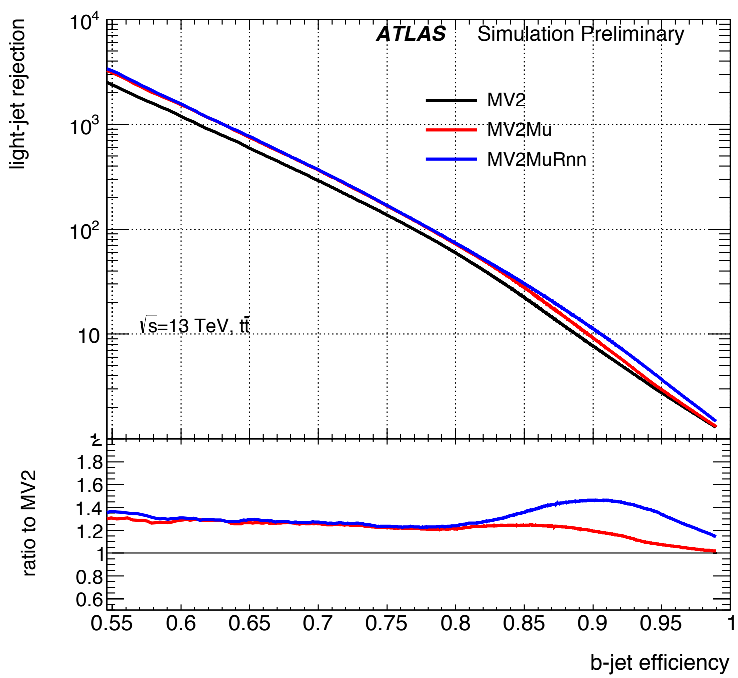




Future Developments

- Deep learning algorithms (DL1)
- Recursive Neural Network (RNN)
  - Takes into account track IP correlations to improve performance
- Dedicated c-tagging algorithms (H→cc):
  - Use DL algorithms or 2D cut on MV2c100 and MV2cl100
  - More limited efficiency than b-tagging (typical WP ~ 25% c-jet efficiency)
- Inclusion of muon variables:
  - 10% of b-hadron decays produce a muon





Useful links

» b-jet performance note (2017) - https://cds.cern.ch/record/2273281

» b-jet performance note (2016) - https://cds.cern.ch/record/2160731

» b-jet performance note (2015) - https://cds.cern.ch/record/2037697

» b-jet efficiency measurement (2012) - https://cds.cern.ch/record/1664335

» ATLAS Flavour Tagging Public Page - https://twiki.cern.ch/twiki/bin/view/AtlasPublic/FlavourTaggingPublicResultsCollisionData

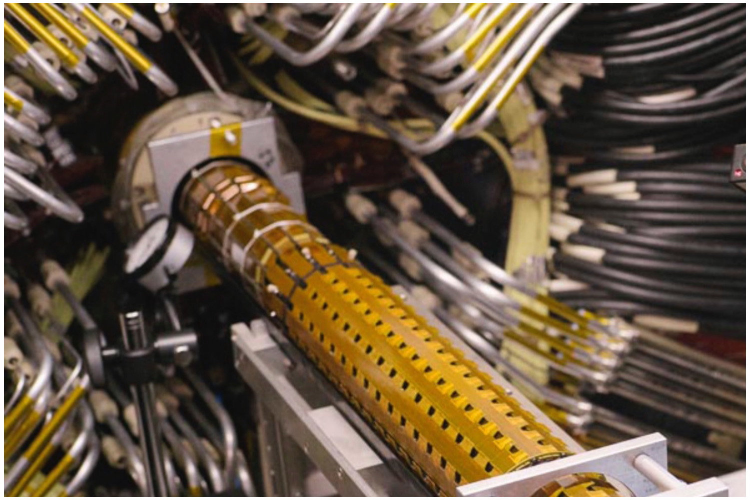
» Comparison of MC generator predictions for b- and c- hadrons in the decays of top quarks and the fragmentation of high p_T jets:

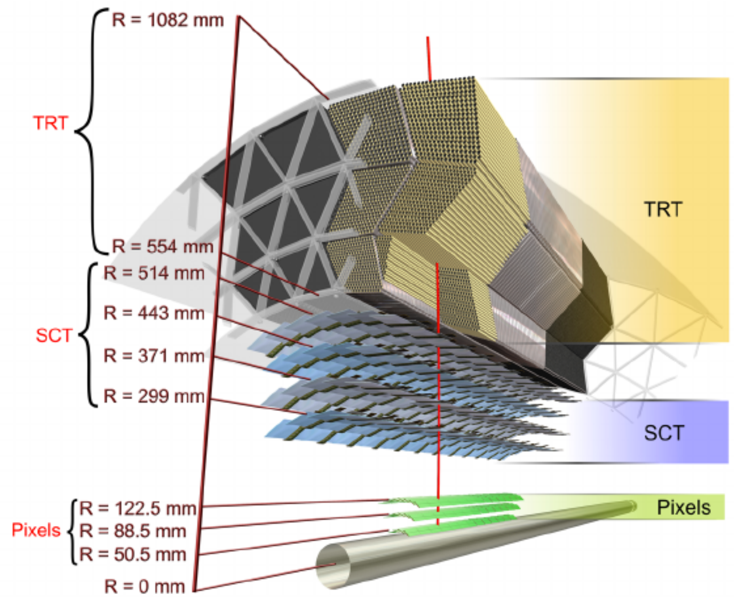
  » https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2014-008/

Back-Up

Tracking in ATLAS

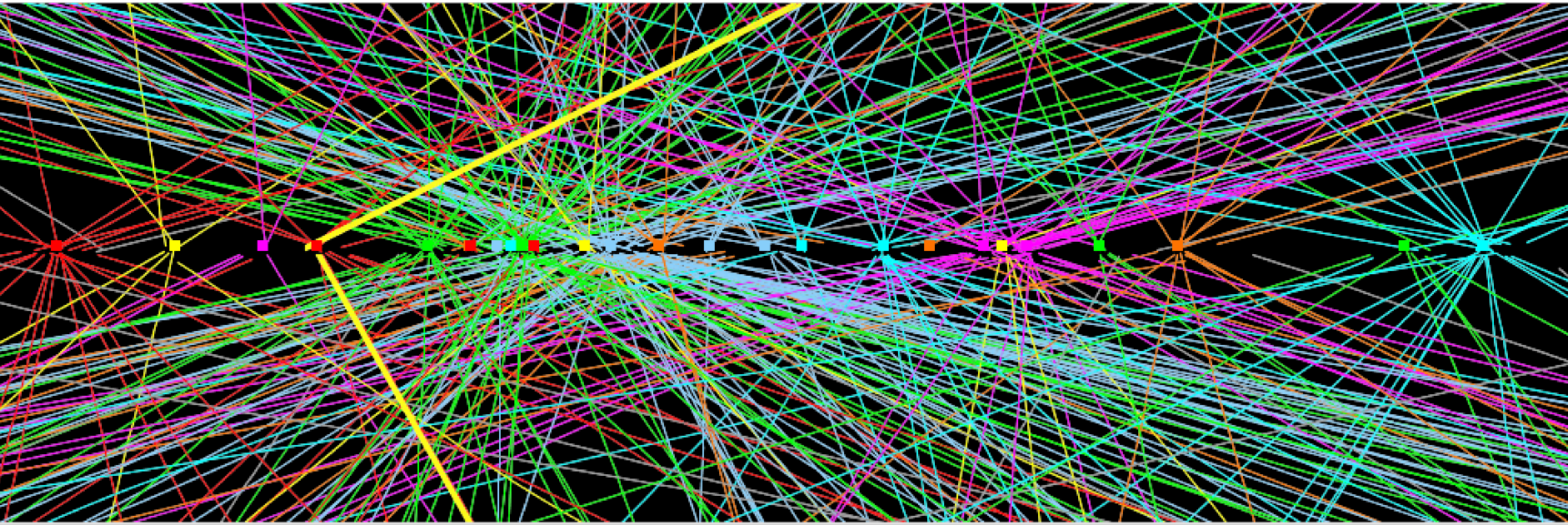
» Tracks used for b-tagging are reconstructed using the ATLAS inner detector (ID)
  » Immersed in a 2 T solenoid magnetic field (essential for track momentum measurement)
  » Insertable B-Layer (**new for Run-2**) is the inner-most layer at ~30mm radius (silicon pixels)
    » Greatly improved track resolution over Run-1
  » Next layers (radially outwards): **pixel layers**, **semiconductor tracker** and **transition radiation tracker**
» Tracks of charged particles are reconstructed using hits in the ID
» Associated to jets using a ΔR(track, jet) matching (jet p_T dependant)



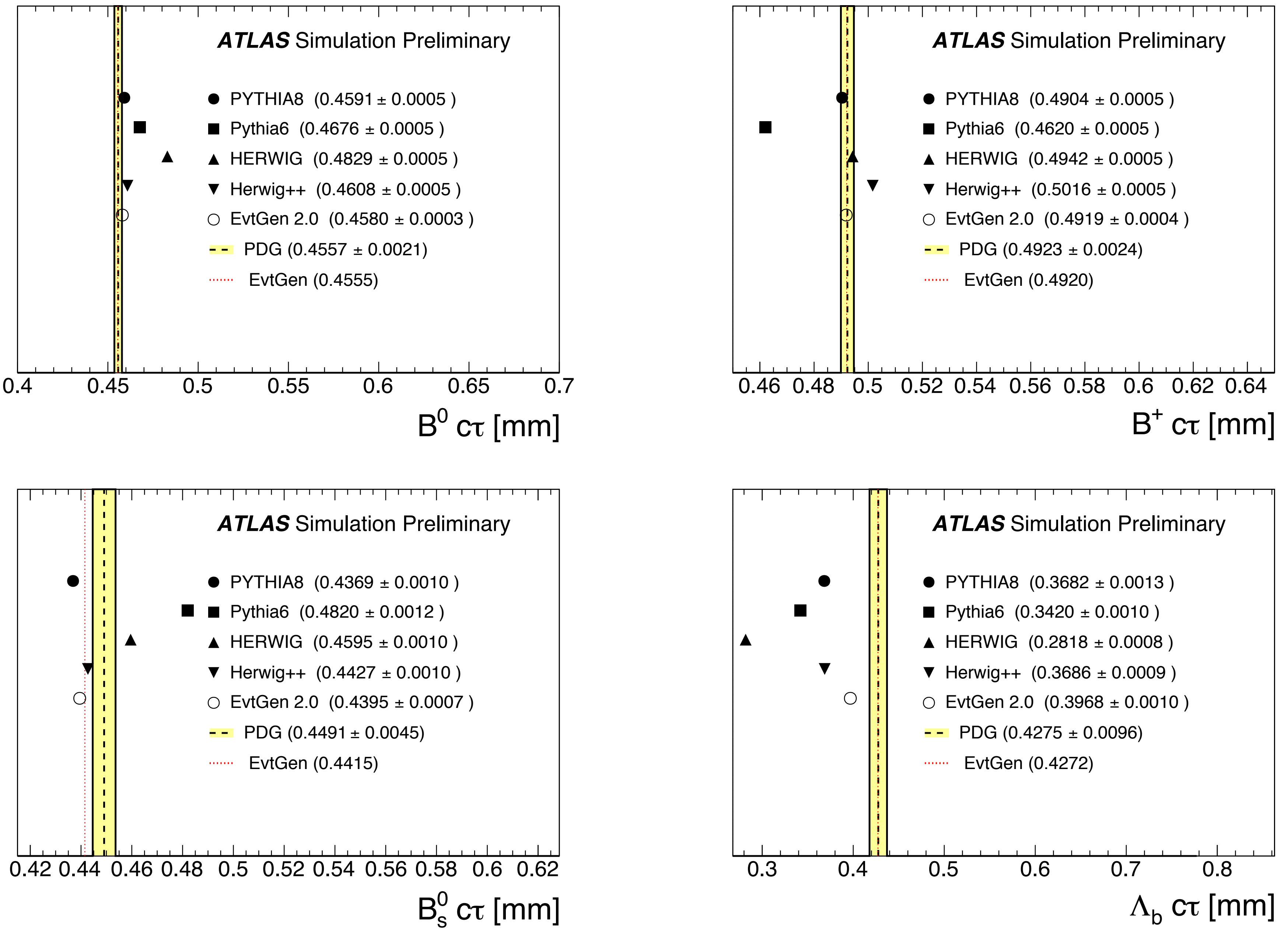


Primary Vertex Reconstruction

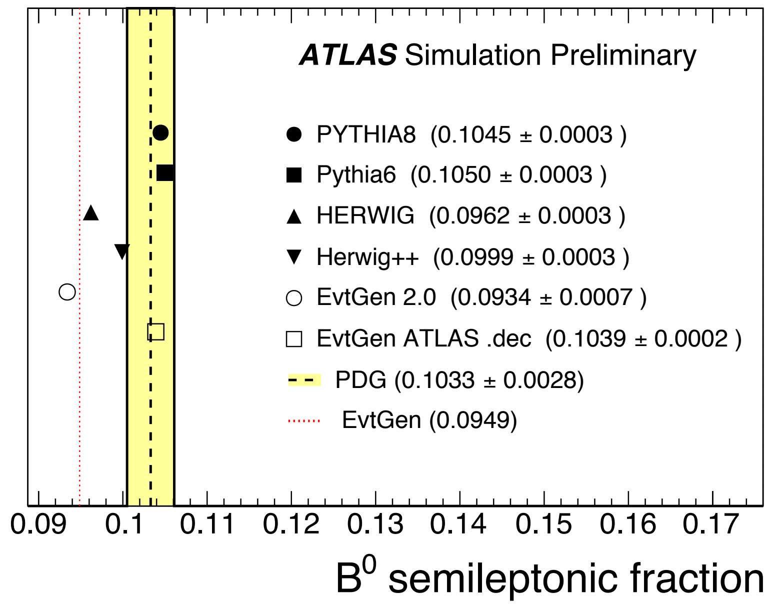
- Colliding bunches contain up to 10^{11} protons
  - In 2016, up to 40 interactions per bunch crossing
  - Need to effectively find the primary vertex of the event of interest
- Reconstruct all vertices using ID tracks (example below has ~25 vertices) and an iterative procedure
  - Select vertex seed from beamspot and track information
  - Remove tracks from seed until passes set quality requirement
  - Use discarded tracks to seed another vertex, and repeat until all tracks associated to a vertex
- Primary vertex is chosen as the one with highest sum of squared track p_T
  - Correctly identifies the primary vertex ~99% of the time

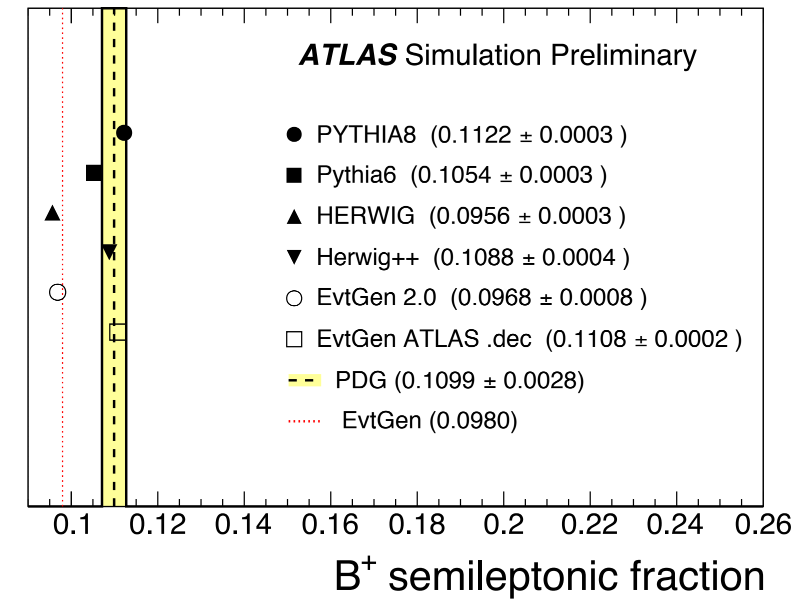


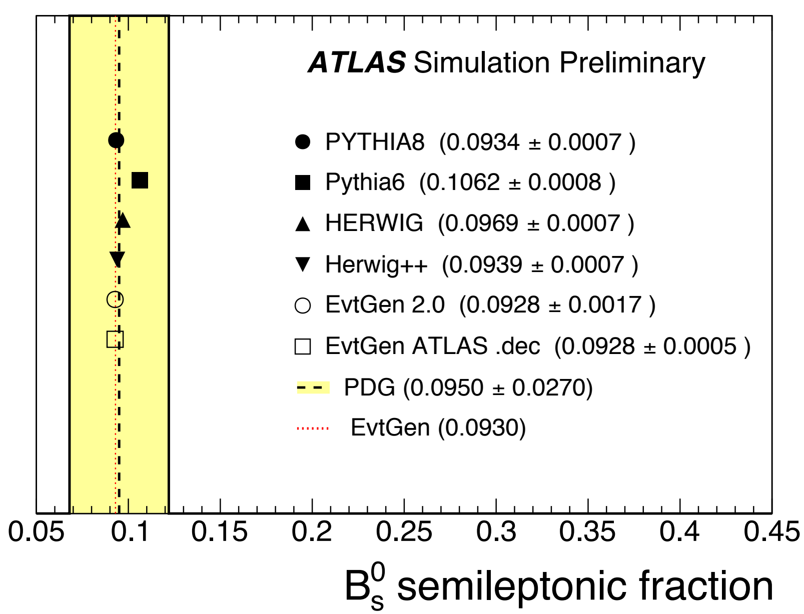
b-Hadron Lifetimes

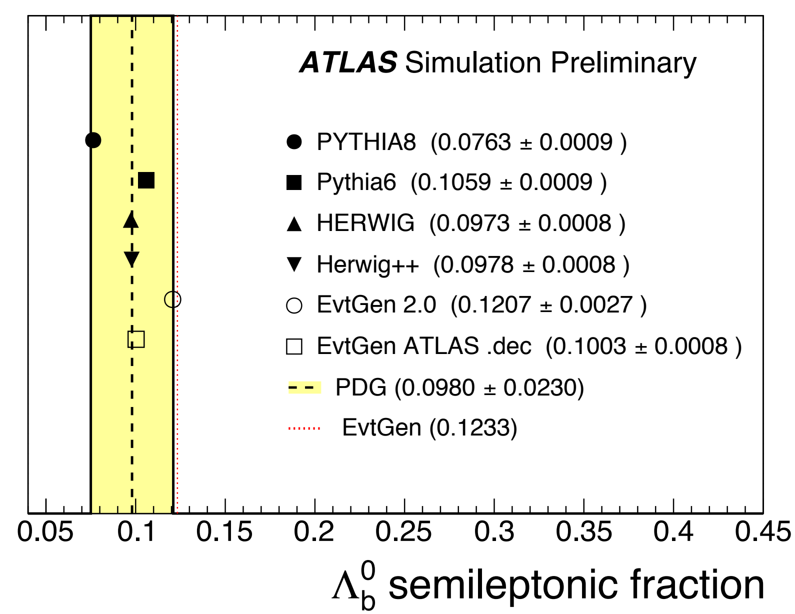


b-Hadron Semileptonic Fractions

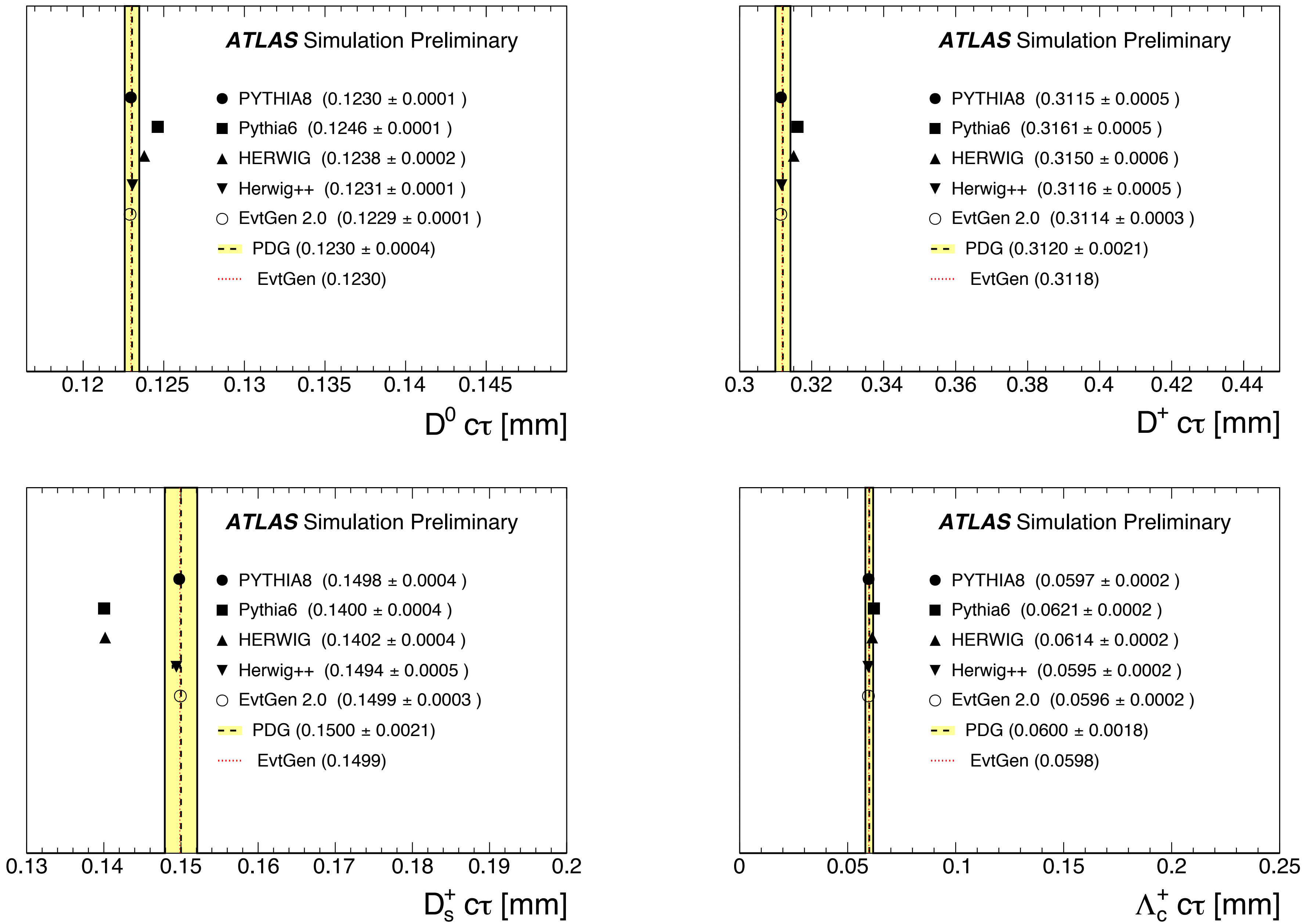








c-Hadron Lifetimes



b-Hadron Semileptonic fractions