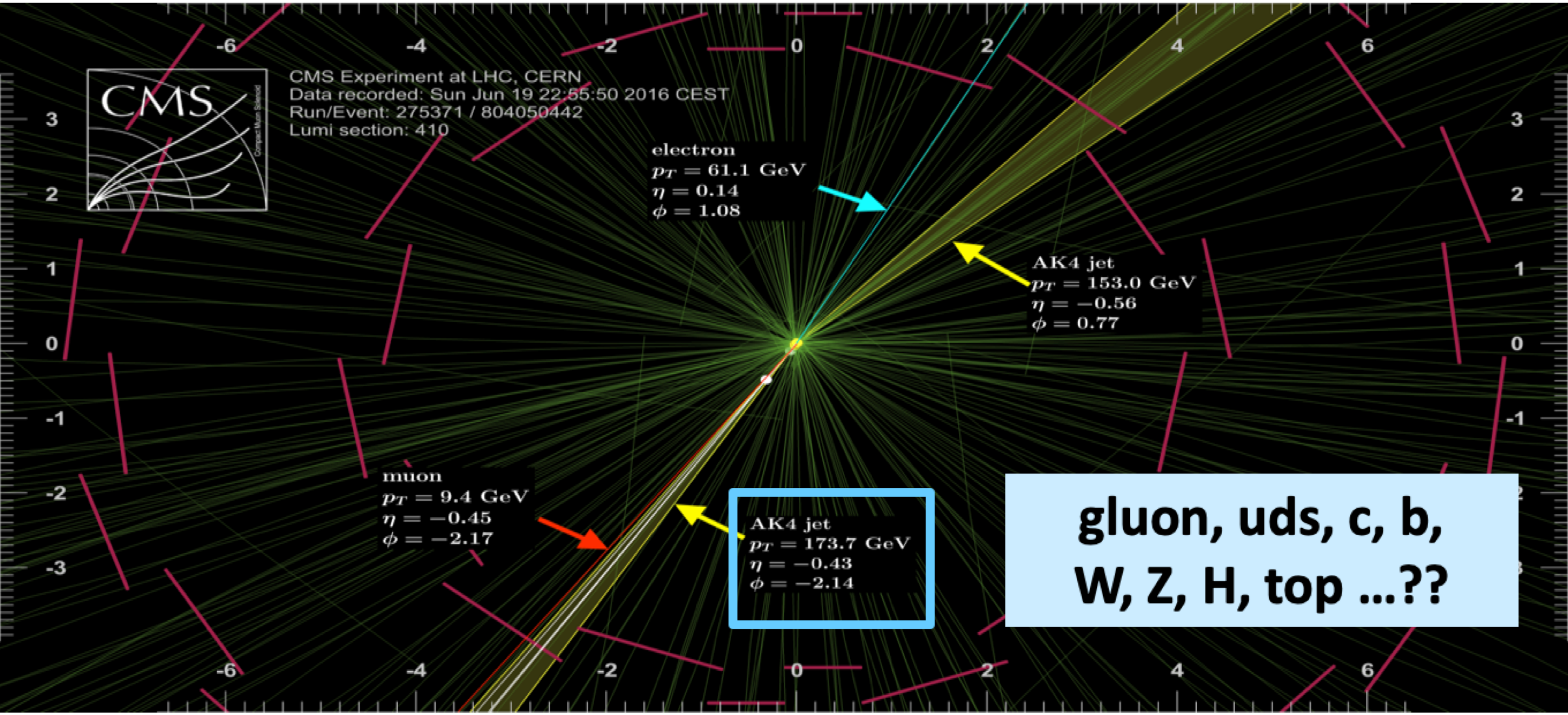
Jet Flavour Tagging at the FCC-ee

Franco Bedeschi, Loukas Gouskos,
Michele Selvaggi

RD_FCC Collaboration meeting
15-16 Dec, 2021

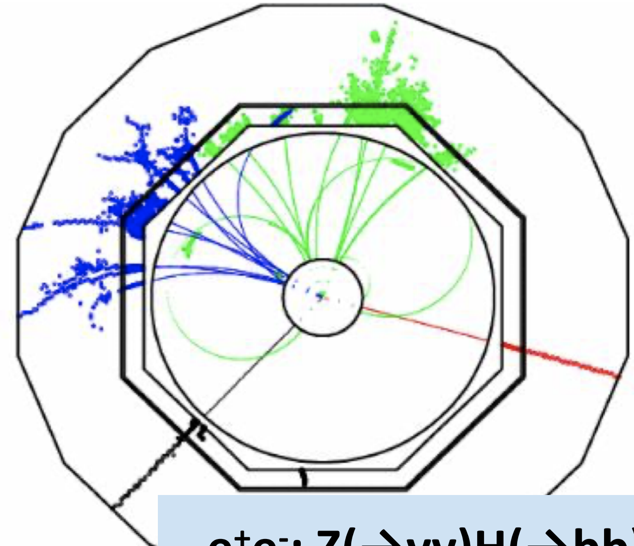
Introduction: Jet tagging



- Jet tagging: almost 30 years at colliders
  - b jets at LEP & Tevatron, then top, W/Z and Higgs jets at the LHC
- Recently: start developing powerful and multi-object tagging capabilities
  - potential to open access to many new physics topics that had been written off previously

Main goals

- e+e-: colliders provide a very clean environment
  - Much lower occupancy, no pileup compared to LHC

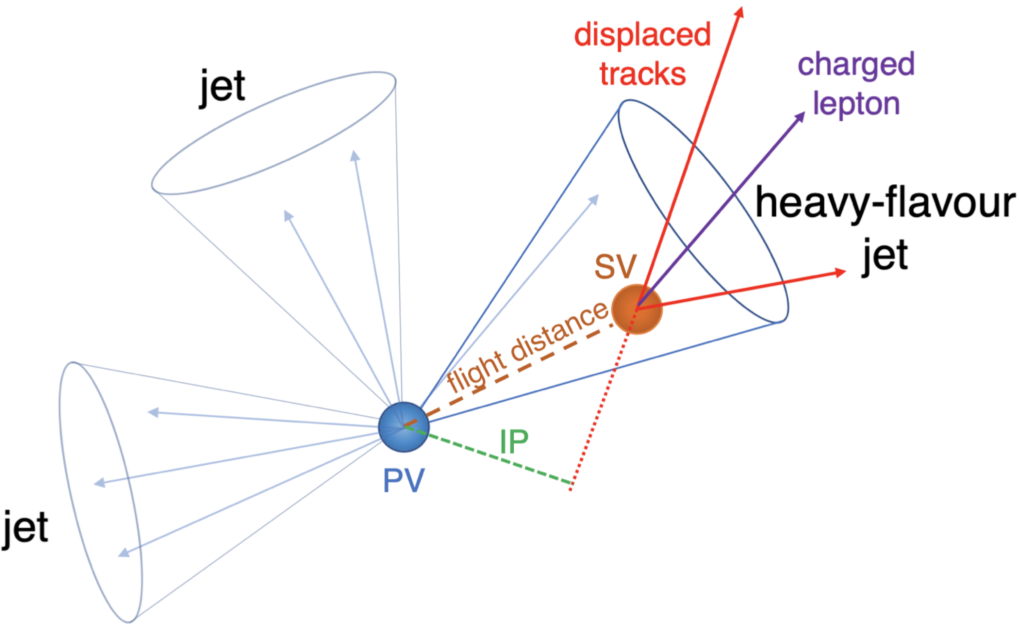


e^+e^-: Z(→νν)H(→bb)

- Understand detector requirements/optimize design
  - Vertexing and PID capabilities of the FCCee detectors

- Develop a versatile jet flavor tagger for FCCee
  - Identify with high purity gluon / light / strange / charm / bottom quarks
    - multi-class classifier

Basics of flavour tagging (b/c)

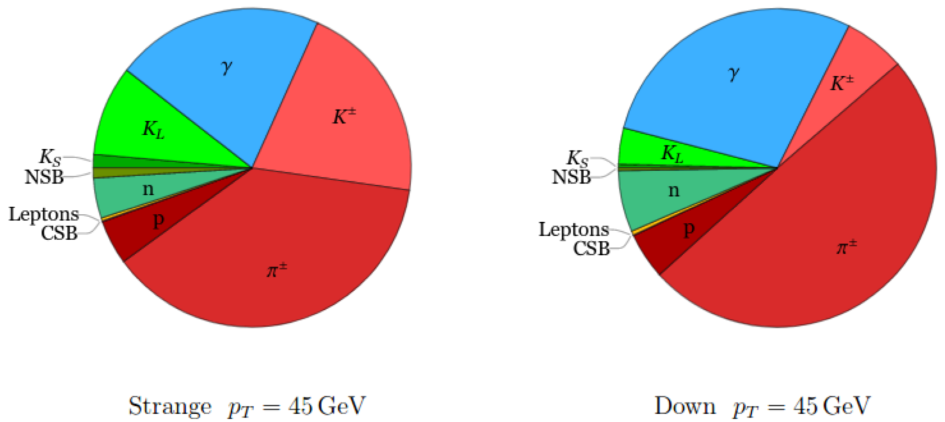


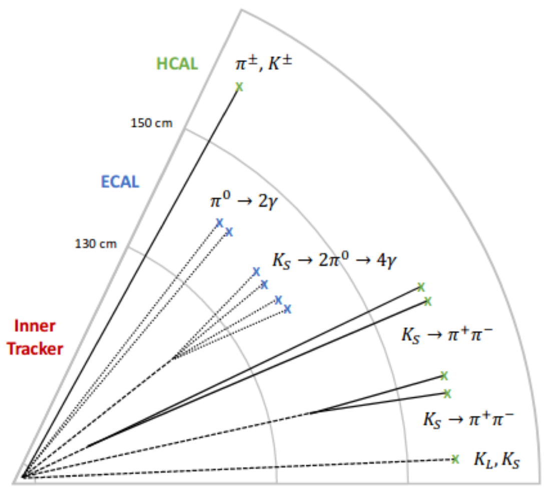
Detector constraints:
Need power pixel/tracking detectors
- Good spatial resolution
- As little material as possible
- Precise track alignment

- Large lifetime
  - b (c) lifetime ~ps (~0.1ps)
  - b (c) decay length: ~5 (2-3) mm for ~50 GeV boost
- Displaced vertices/tracks
  - Large impact parameters
  - Tertiary vertices when B hadron decays to C hadron
- Large track multiplicity
  - ~5 (~2) charged tracks/decay
- Non-isolated e/μ
  - ~20 (10)% in B (C) decays

Basics of flavour tagging (strange)

[2003.09517] Momentum weighted fraction:





- Large Kaon content
  - Charged Kaon as track:
    - K/pi separation
      - TOF
      - dEdx/dNdx
  - Neutral Kaons:
    - $K_S \to \pi\pi$
      - Displaced 2 track vertex
      - 4 photons
    - K_L
      - TOF vs n ?

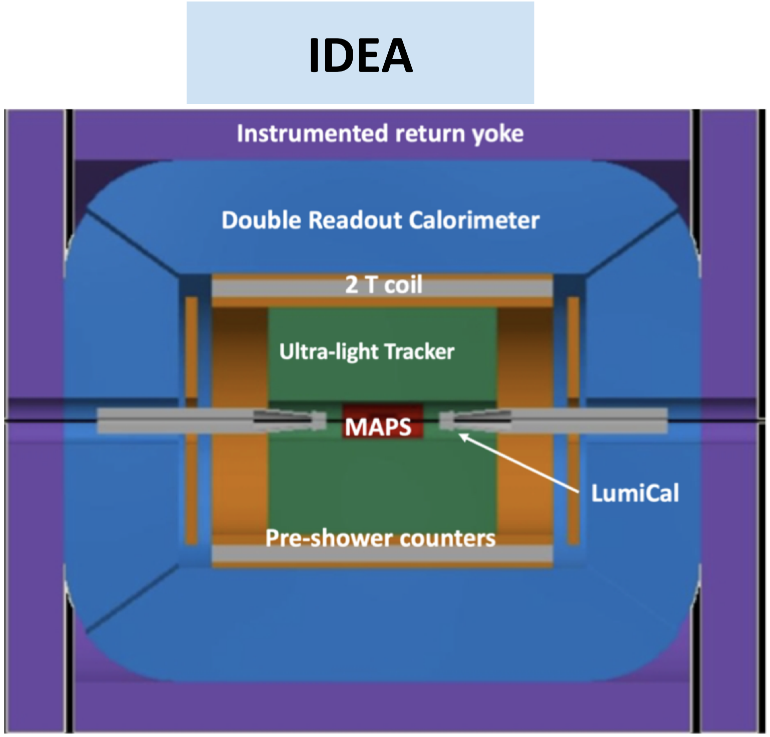
Detector constraints:

Need power pixel/tracking detectors

- good spatial resolution
- timing detectors
- charged energy loss (gas/silicon)

FCCee detector

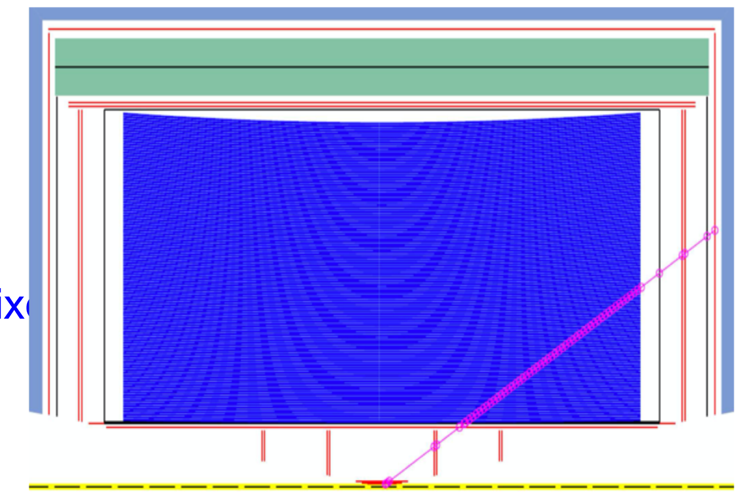
- Ideal for flavour identification [hence: measure Higgs couplings]
  - Impact parameter resolution
    - Low material budget tracker (minimise multiple scattering)
    - Small beam-pipe 1.5 cm -- investigating 1 cm
  - PID capabilities
    - dEdx (Si tracker) -- Cluster counting (Drift)
    - Time of flight -- timing layer



Simulation

- Detector response based on Delphes:
  - Including FastTrackCovariance
  - Computes:
    - full track covariance matrix (5x5)
      - Including MS
    - smeared track using the off-diagonal terms
    - path length and dN/dx for various gas mix
      - Allows fast turn-around when trying different detector options

IDEA



- MC Samples:
  - MG5+Pythia8 used to generate:
    - ee→ ZH→vvXX events (X: g, ud, s, c, b)

- Jets clusters with the generalized-kT algorithm using *R*=1.5
  - Similar to the anti-k_T algorithm [IRC safe]

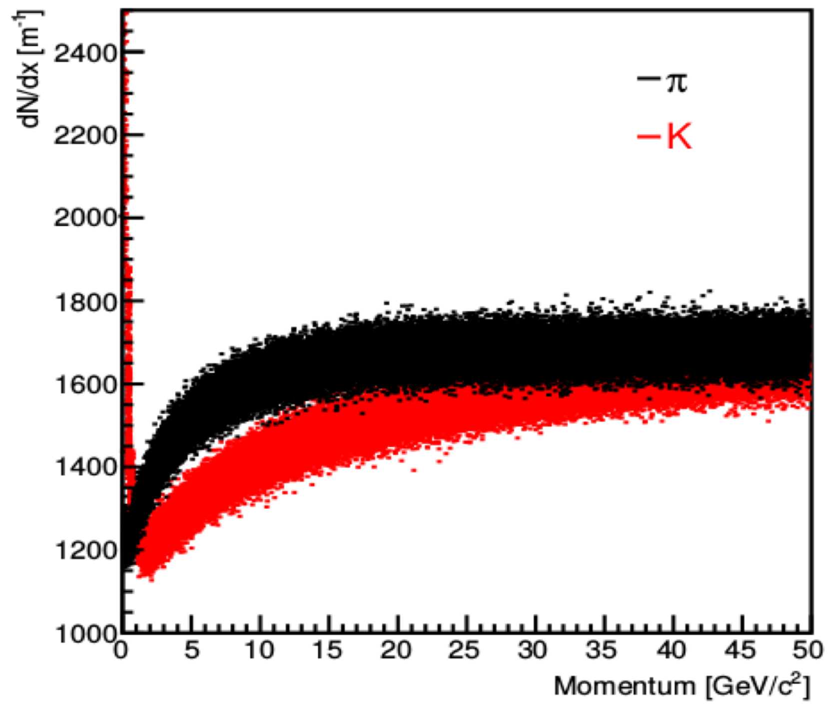
Cluster counting dN/dx

- Count number of **primary ionisation** clusters along track path
- Avoids large landau flukes (**poisson distributed**)
- Requires high granularity
- Module added in Delphes

```
###################
# Cluster Counting
###################

module ClusterCounting ClusterCounting {

  add InputArray TrackSmearing/tracks
  set OutputArray tracks

  set Bz $B

  ## check that these are consistent with DCHCANI/DCHNANO parameters in TrackCovariance module
  set Rmin $DCHRMIN
  set Rmax $DCHRMAX
  set Zmin $DCHZMIN
  set Zmax $DCHZMAX

  # gas mix option:
  # 0:  Helium 90% - Isobutane 10%
  # 1:  Helium 100%
  # 2:  Argon 50% - Ethane 50%
  # 3:  Argon 100%

  set GasOption 0

}
```

IDEA detector:

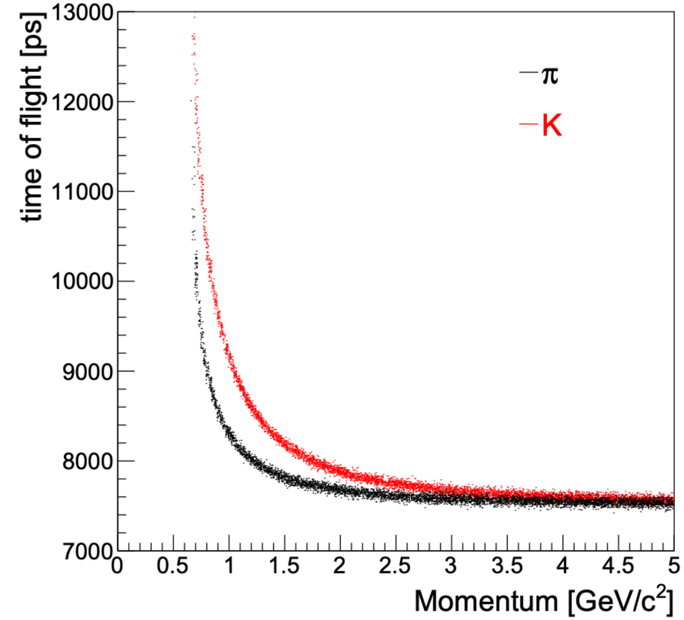
90% He / 10 % Isobutane



Time-of-flight

- Allows for good K/pi separation at low momenta:

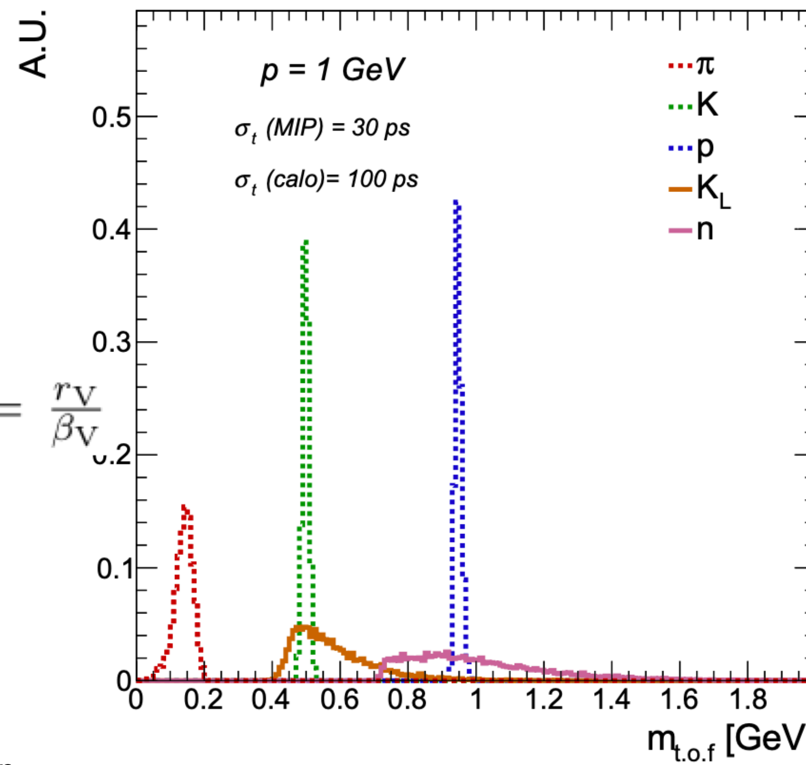
$$t_{\text{flight}} \equiv t_{\text{F}} - t_{\text{V}} = \frac{L}{\beta} = \frac{L\sqrt{p^2 + m^2}}{p}$$



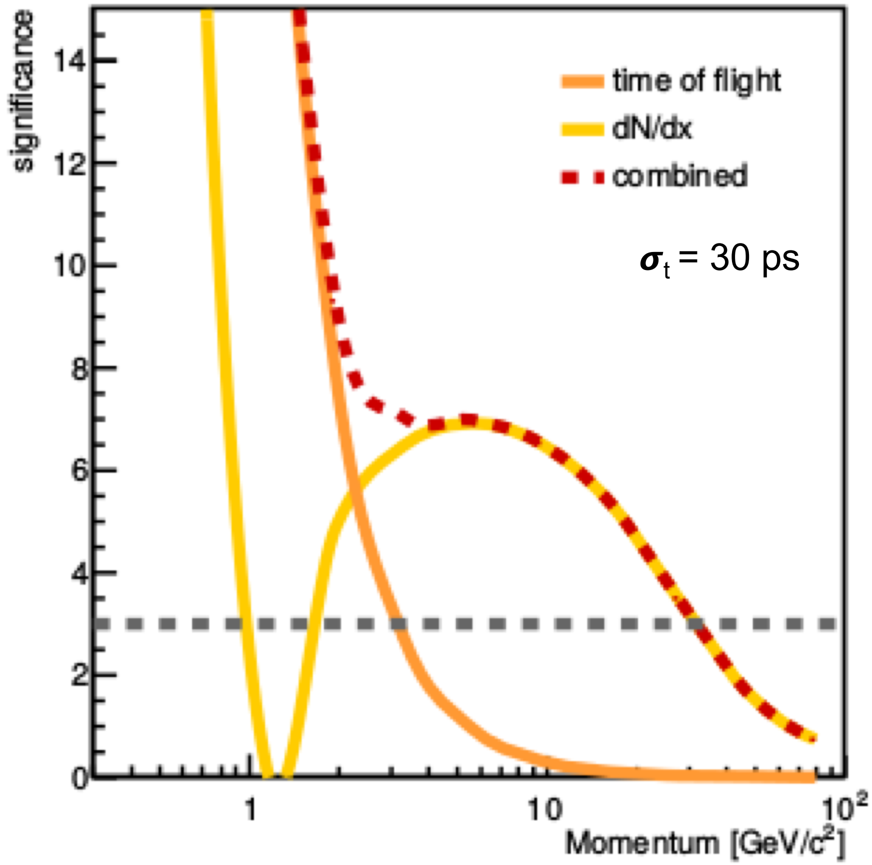
- Need to make assumption on vertex time (crucial for highly displaced K_S) :

```
########################################
#   Time Of Flight Measurement
########################################

module TimeOfFlight TimeOfFlight {
  set TrackInputArray TimeSmearing/tracks
  set VertexInputArray TruthVertexFinder/vertices

  set OutputArray tracks

  # 0: assume vertex time tV from MC Truth (ideal case)
  # 1: assume vertex time tV = 0
  # 2: calculate vertex time as vertex TOF, assuming tPV=0

  set VertexTimeMode 2
}
```

$$t_{\text{V}} = \frac{r_{\text{V}}}{\beta_{\text{V}}}$$



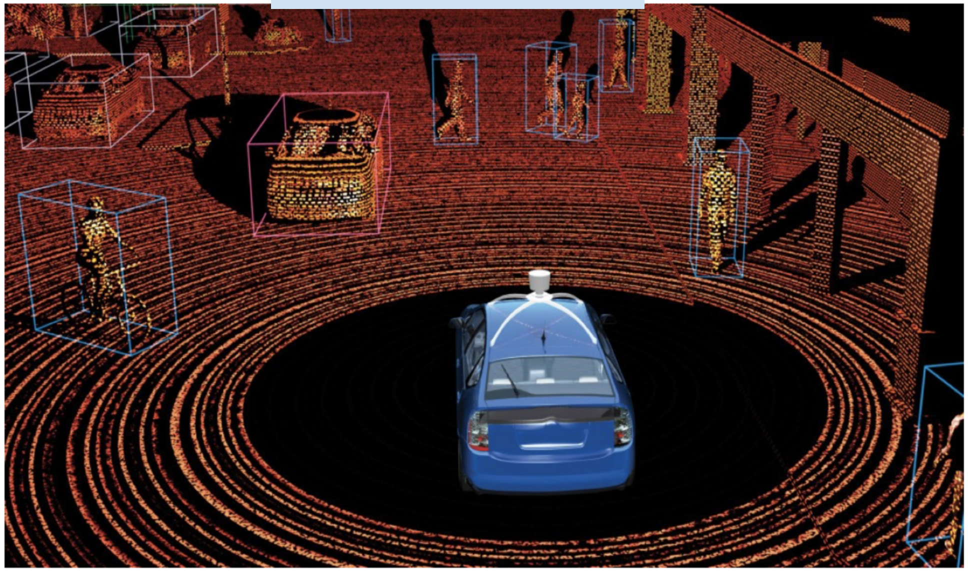
Combined PID



3 std deviation K/pi separation for tracks with p < 30 GeV

Designing a jet flavour tagging algorithm

A point cloud

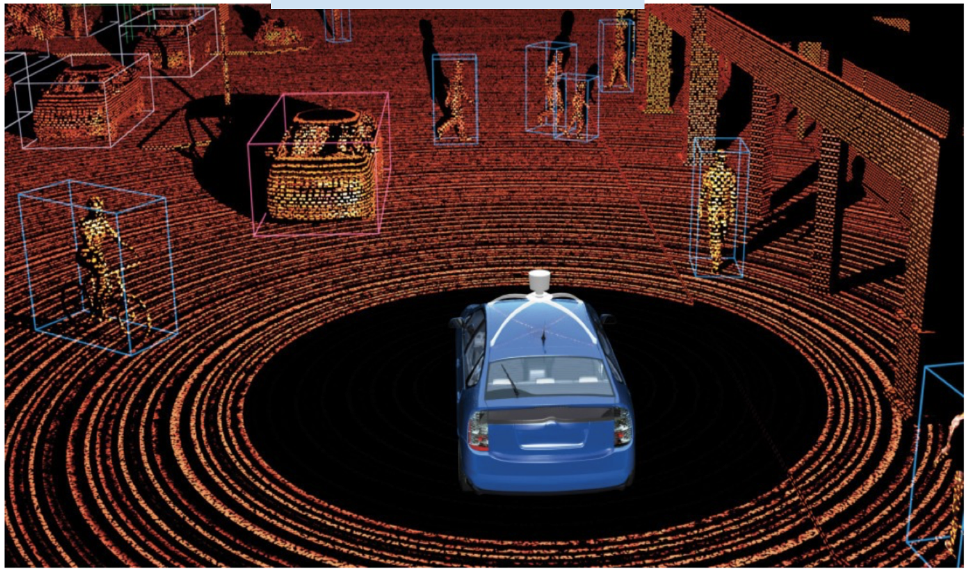


Source:https://news.voyage.auto/an-introduction-to-lidar-the-key-self-driving-car-sensor-a7e405590cff

- Point cloud (Wikipedia):
  - A set of data **points** in space
  - Produced by 3D scanners, which measure a large number of points on the external surfaces of objects around them

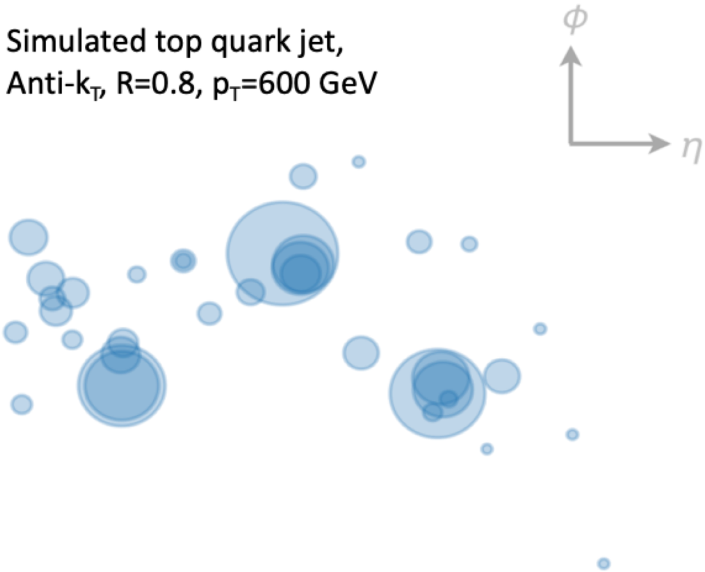
From point clouds to particle clouds

A point cloud



Source: https://news.voyage.auto/an-introduction-to-lidar-the-key-self-driving-car-sensor-a7e405590cff

A particle cloud



- Point cloud (Wikipedia):
  - A set of data **points** in space
  - Produced by 3D scanners, which measure a large number of points on the external surfaces of objects around them
- Particle cloud :
  - A set of **particles** in space
  - Produced by clustering a large number of particles measured by the detectors

Flavour tagging using ParticleNet

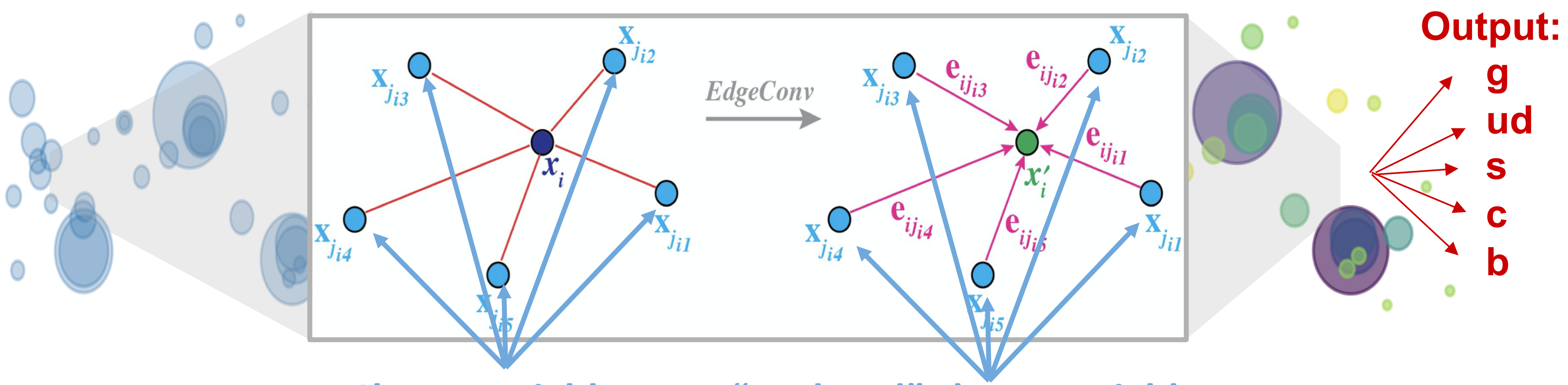
[H. QU, LG](#)
[PRD 101 (2020) 5, 056019](#)
[CMS-DP-2020-002](#)

- Developing a flavour tagging algorithm based on ParticleNet
  - Jet is represented as a “particle cloud”
- Follow a hierarchical learning approach:
  - **First:** Learn “local” structures; **Then:** move to more “global” features
  - Treat the particle cloud as a graph
    - **Particles** are the **vertices** of the graph

      Relationships between the particles are the **edges** of the graph

Jet:
As particle cloud

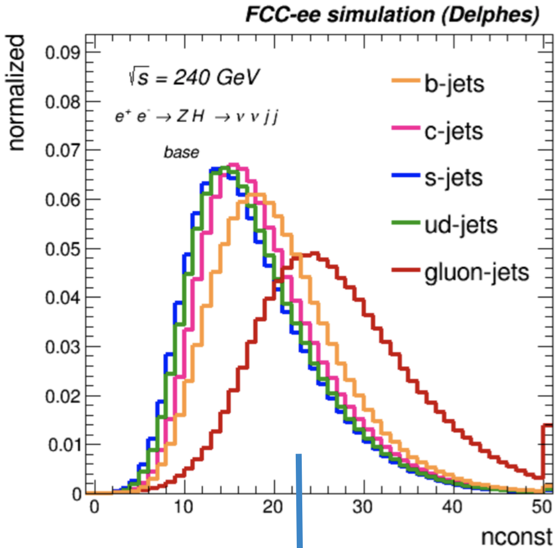
Identify “neighboring” particles



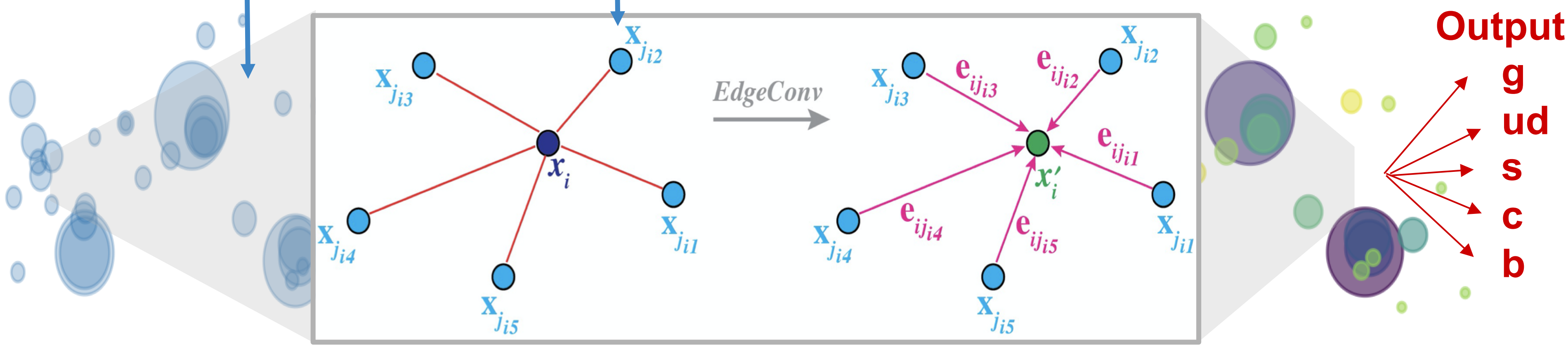
Closest neighbors
[in θ-ϕ]

“Updated” closest neighbors
[from learned features]

Flavour tagging using ParticleNet (II)



Particle features: O(20)/particle — **Particle kinematics, particle charge, Impact parameter (d0, dz) and significance, particle type (el, mu, γ,...)**



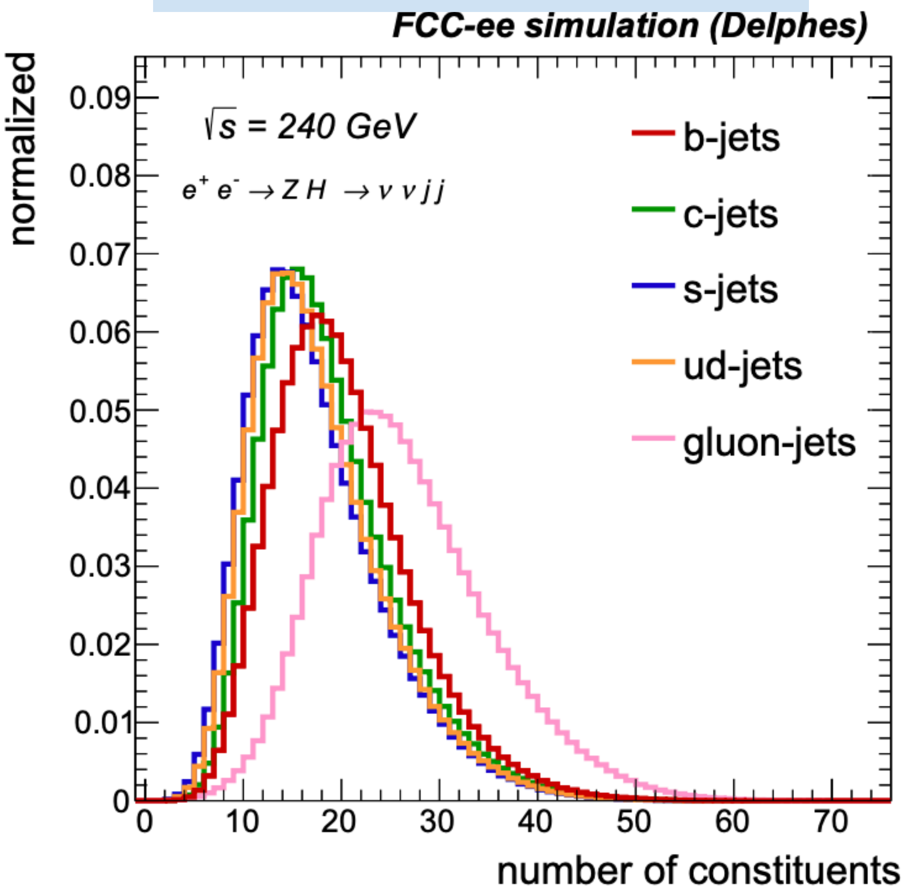
Inputs:
75 particles/jet

Training details:
- 1M jets split equally between classes
- Lots of room for improving the training details

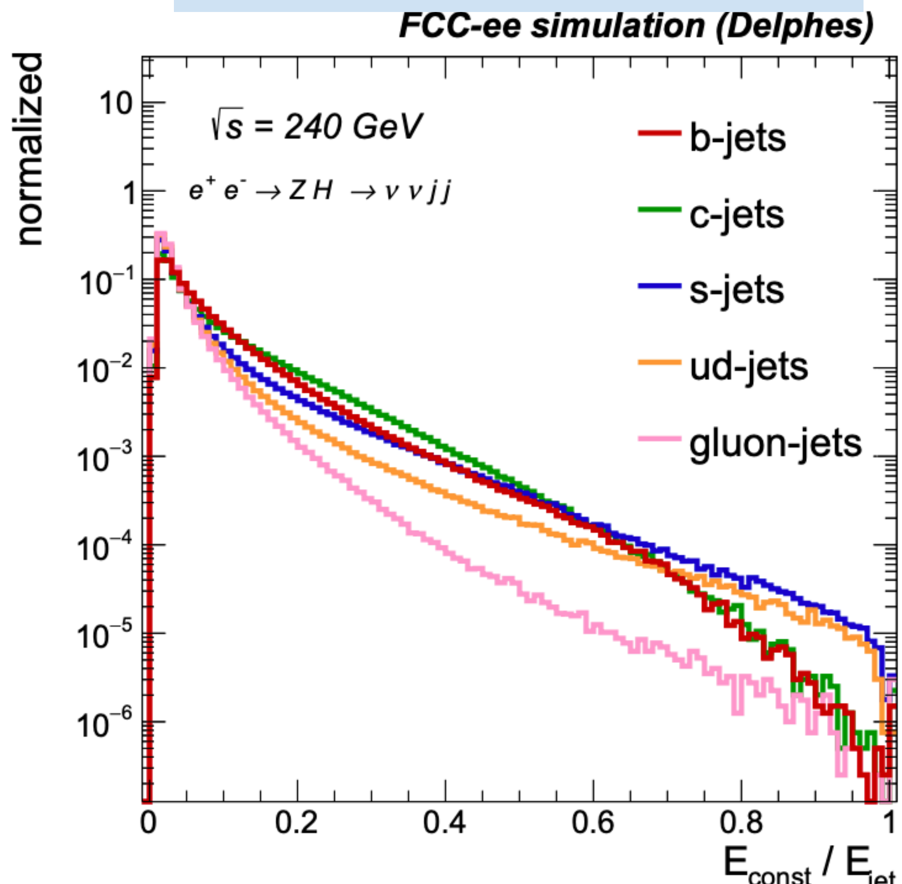
Input variables

- Comparison of input distributions for different jet flavors

Number of jet constituents



Constituent relative energy

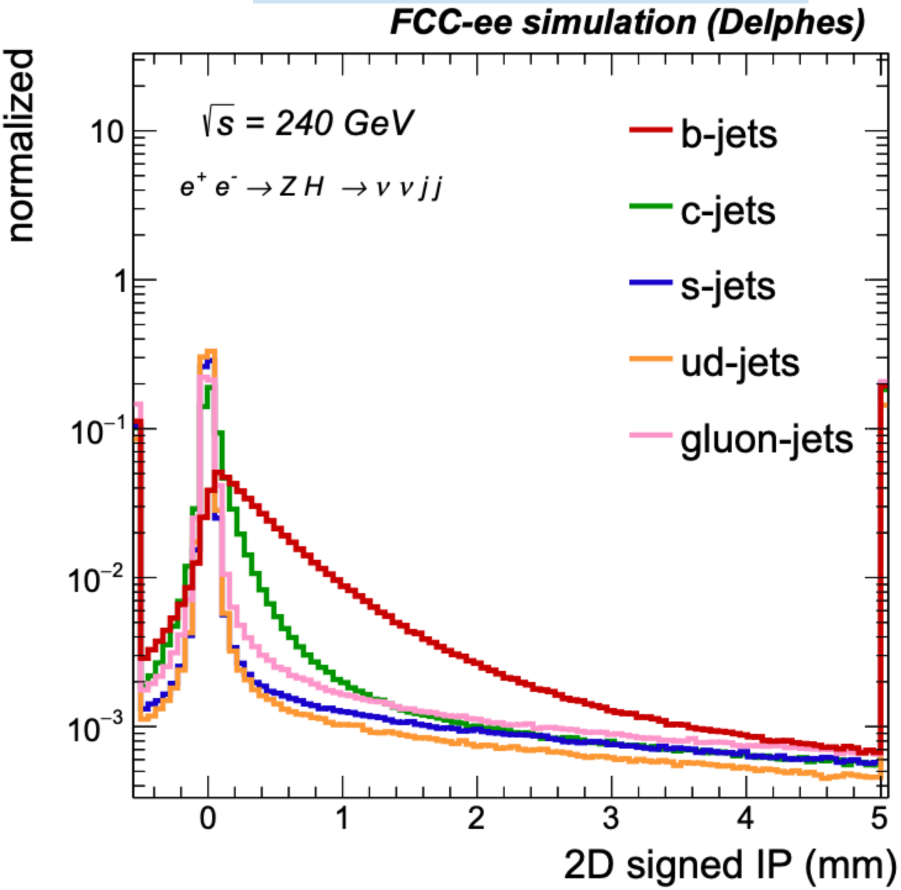


- More comparisons: https://selvaggi.web.cern.ch/selvaggi/FCC/FCCee/FlavourTagging/

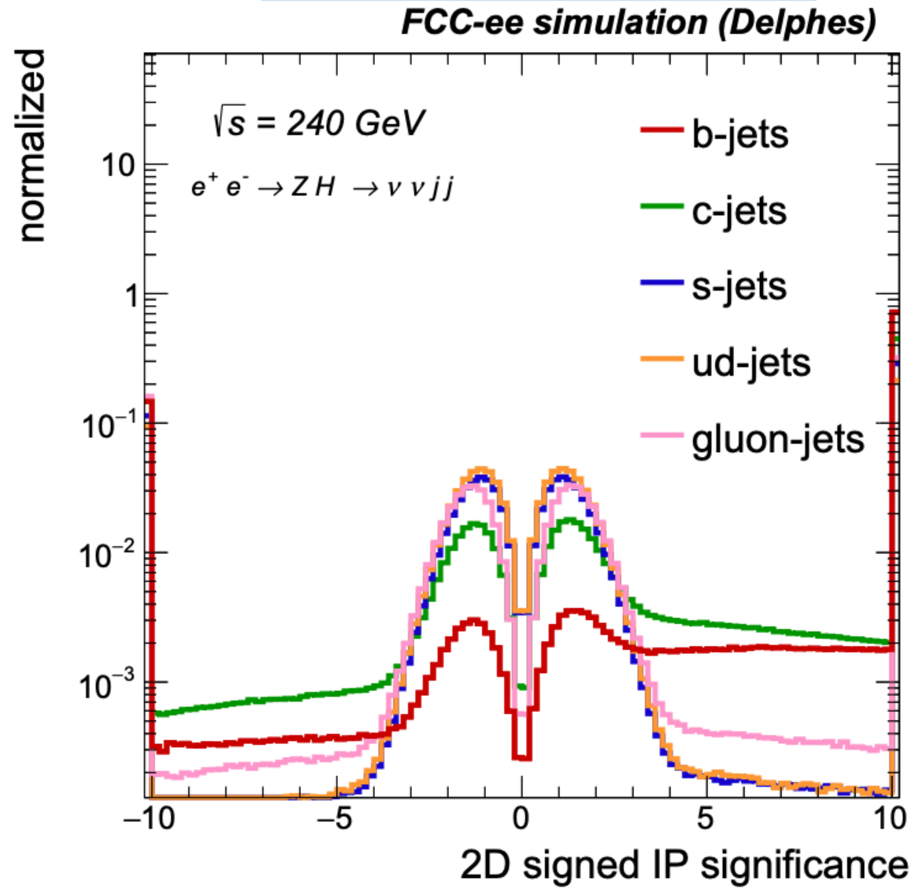
Input variables

- Comparison of input distributions for different jet flavors

Impact parameter (d_0)



d_0 significance



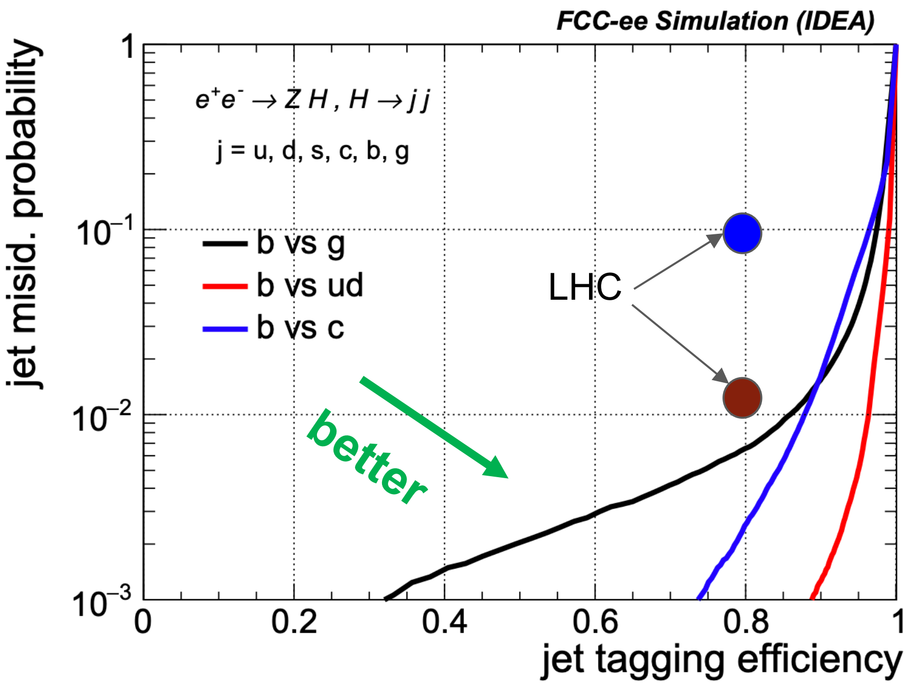
- More comparisons: https://selvaggi.web.cern.ch/selvaggi/FCC/FCCee/FlavourTagging/

Full list of input variables

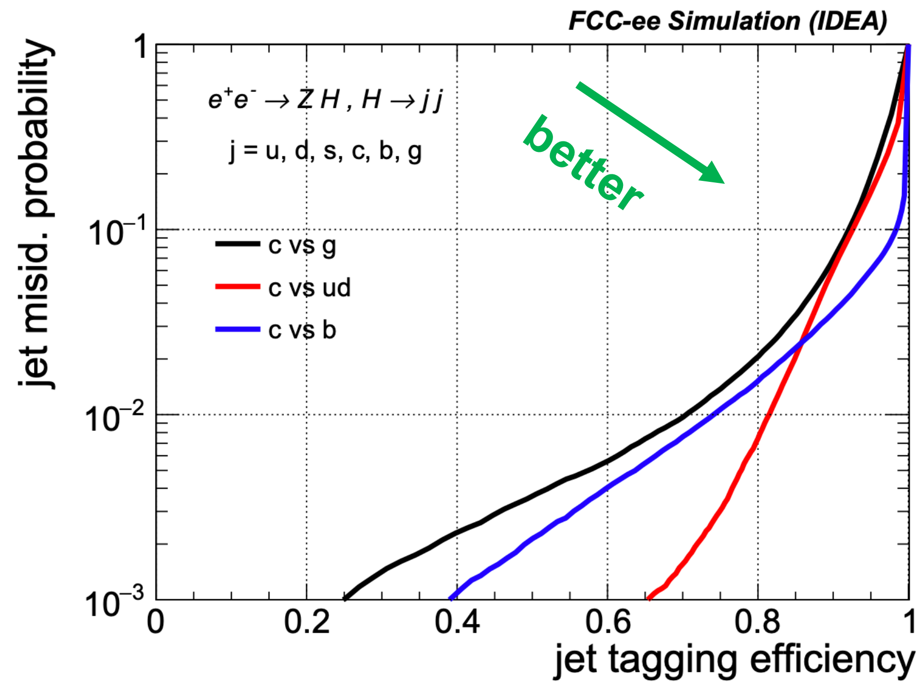
Variable	Description
	Kinematics
$E_{\text{const}}/E_{\text{jet}}$	energy of the jet constituent divided by the jet energy
$\sin(\theta_{\text{jet,const}})$	sin of the angle between the constituent momentum and the jet momentum
$\cos(\theta_{\text{jet,const}})$	cos of the angle between the constituent momentum the jet momentum
	Displacement
SIP_{2D}	signed 2D impact parameter of the track
$\text{SIP}_{\text{2D}}/\sigma_{\text{2D}}$	signed 2D impact parameter significance of the track
SIP_{3D}	signed 3D impact parameter of the track
$\text{SIP}_{\text{3D}}/\sigma_{\text{3D}}$	signed 3D impact parameter significance of the track
d_{3D}	jet track distance at their point of closest approach
$d_{\text{3D}}/\sigma_{d_{\text{3D}}}$	jet track distance significance at their point of closest approach
	Identification
q	electric charge of the particle
$m_{\text{t.o.f.}}$	mass calculated from time-of-flight
dN/dx	number of primary ionisation clusters along track
isMuon	if the particle is identified as a muon
isElectron	if the particle is identified as an electron
isPhoton	if the particle is identified as a photon
isChargedHadron	if the particle is identified as a charged hadron
isNeutralHadron	if the particle is identified as a neutral hadron

Performance (b/c)

b-tagging



c-tagging

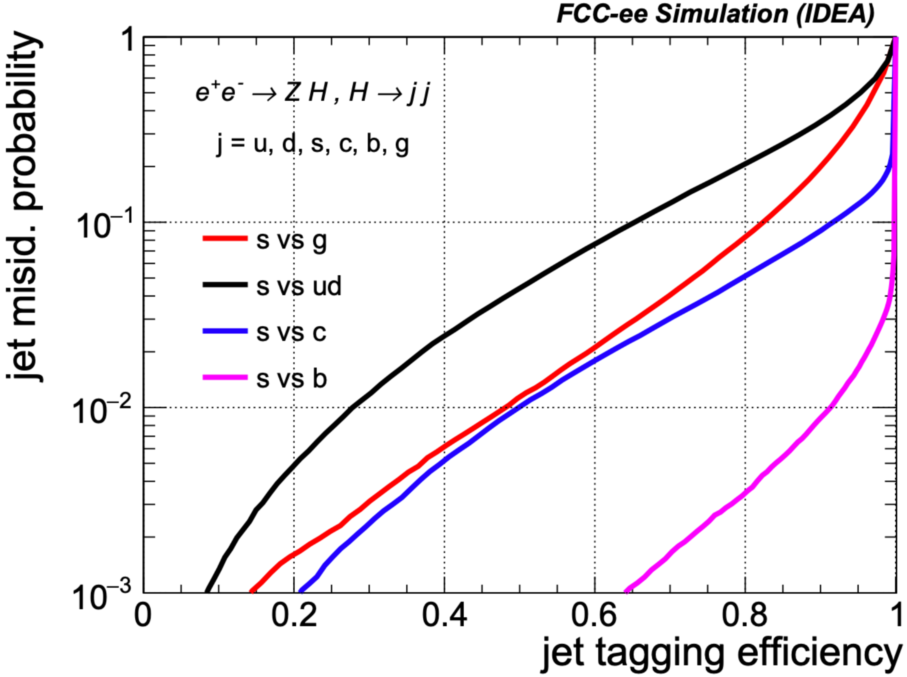


WP	Eff (b)	Mistag (g)	Mistag (ud)	Mistag (c)
Loose	90%	2%	<0.1%	2%
Medium	80%	0.7%	<0.1%	0.3%

WP	Eff (c)	Mistag (g)	Mistag (ud)	Mistag (b)
Loose	90%	8%	8%	4%
Medium	80%	3%	0.7%	2%

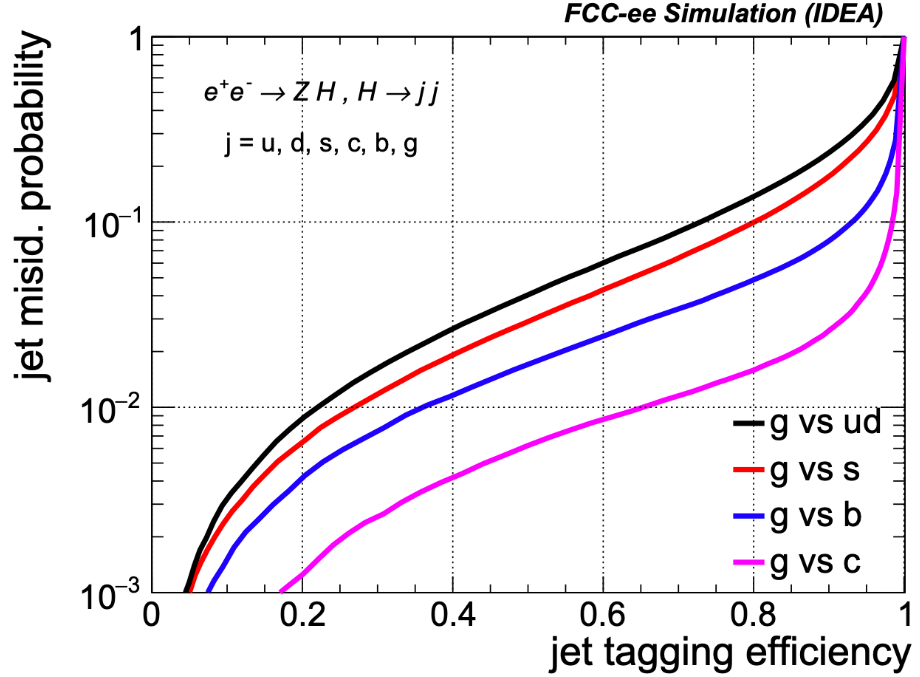
Performance (strange/gluon)

strange-tagging



WP	Eff (s)	Mistag (g)	Mistag (ud)	Mistag (c)	Mistag (b)
Loose	90%	20%	40%	10%	1%
Medium	80%	10%	20%	6%	0.4%

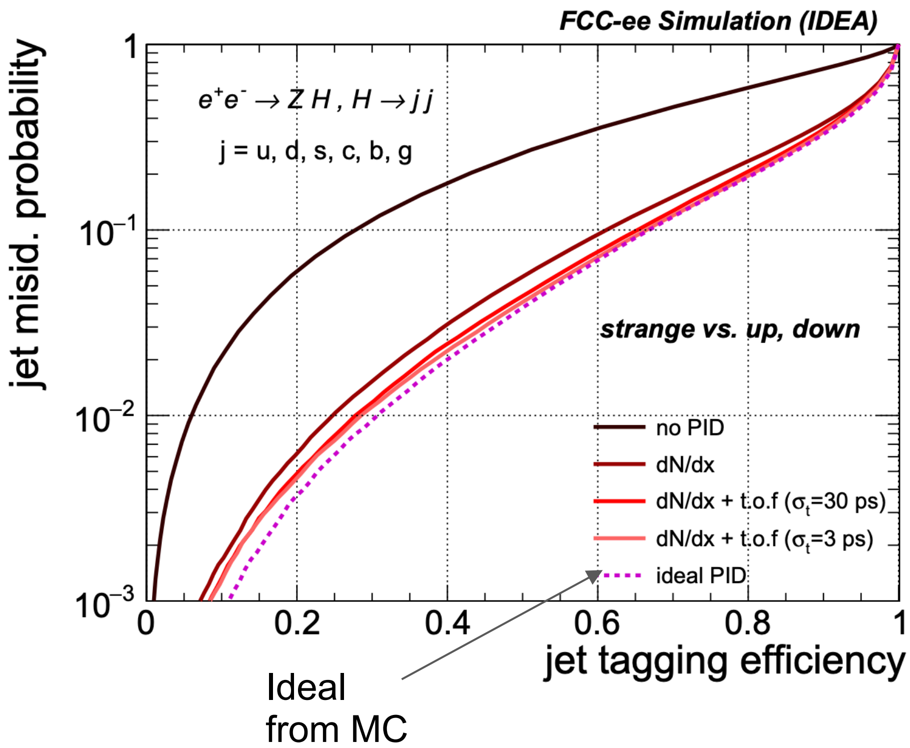
gluon -tagging



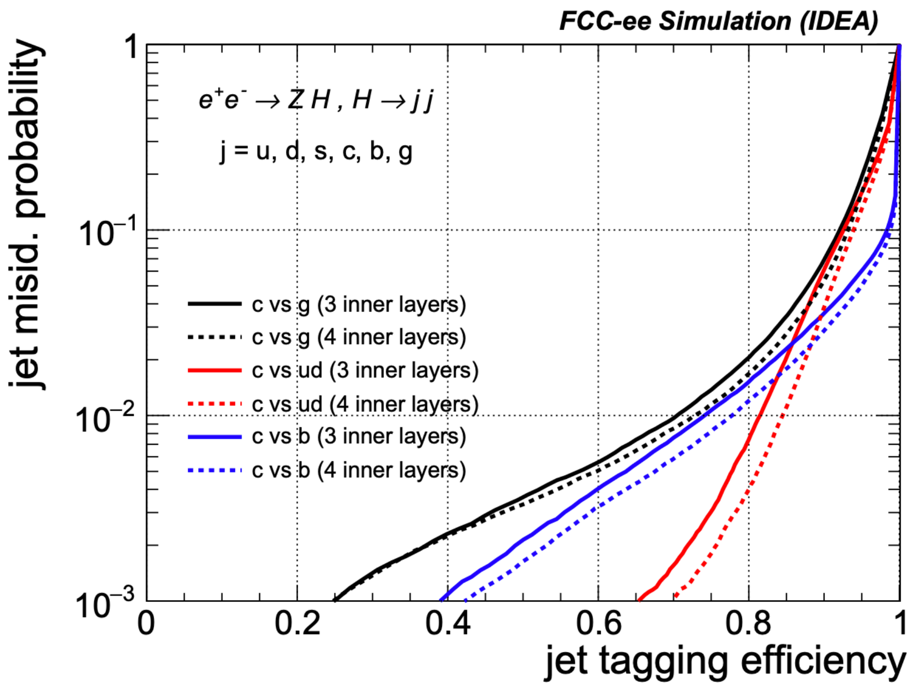
WP	Eff (g)	Mistag (ud)	Mistag (c)	Mistag (b)
Loose	90%	25%	9%	3%
Medium	80%	15%	5%	2%

Impact of detector configurations

Strange tagging [PID]



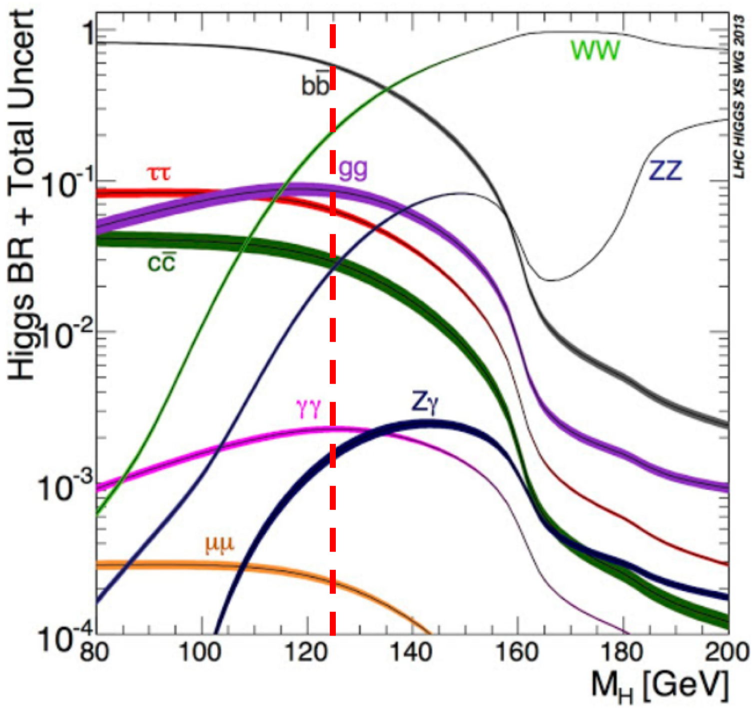
c-tagging [PIX layers]



- Small room for improvement on the PID, in particular for strange tagging
  - TOF does not contribute as much as dNdx (30 ps resolution enough?)
    - low pT tracks are not discriminating ?
    - Can be further improved using timing resolution for neutral K_L vs n ?

REQUIRES FURTHER INVESTIGATION

Higgs couplings: H→cc



$\sqrt{s}$ (GeV)	240		365	
Luminosity (ab^{-1})	5		1.5	
$\delta(\sigma BR)/\sigma BR$ (%)	HZ	$\nu\bar{\nu}$ H	HZ	$\nu\bar{\nu}$ H
$H \to$ any	±0.5		±0.9	
$H \to b\bar{b}$	±0.3	±3.1	±0.5	±0.9
$H \to c\bar{c}$	±2.2		±6.5	±10
$H \to gg$	±1.9		±3.5	±4.5
$H \to W^+W^-$	±1.2		±2.6	±3.0
$H \to ZZ$	±4.4		±12	±10
$H \to \tau\tau$	±0.9		±1.8	±8
$H \to \gamma\gamma$	±9.0		±18	±22
$H \to \mu^+\mu^-$	±19		±40	
$H \to$ invis.	< 0.3		< 0.6	

Ref: Patrick's talk at the CDR Symposium; March 2019

FCCee: σ_{ZH}~200fb, L~5 ab^{-1} (2 IP): **~1M ZH**
[600k H→bb, 100k H→gg, **30k H→cc**]

Use Loose WP:
[c-tag: 90%, b-mistag: 5%, g-mistag: 10%

- **Scenario 2:** Z(→vv)H
  δ(σxBR)/σxBR (%) ~ 1.5 [no systematics]
- **Scenario 1:** Z(→all)H
  δ(σxBR)/σxBR (%) ~ 0.7 [no systematics]

- **Stat limit [i.e. no BKG]:**
  δ(σxBR)/σxBR (%) ~0.6%
- **No BKG rejection:**
  δ(σxBR)/σxBR (%) ~2.9%

Results look promising

Higgs couplings: $H \rightarrow ss$

$BR(H\rightarrow ss) = BR(H\rightarrow cc)(m_s/m_c)^2 \sim 2.3\ 10^{-4}$

FCCee: σ_{ZH}~200fb, L ~ 5 ab^{-1} (2 IP): **~1M ZH**
[600k H→bb, 100k H→gg, 30k H→cc, **200 H→ ss**]

Use Tight WP:
[**s-tag: 60%**, **g-mistag, c-mistagm and b-mist**: negligible

- The most challenging BKG is ZZ
with one Z of shell ~125 GeV [~10% of the Higgs signal]

- Optimistic assumption:
- 100% of the Higgs events (i.e. the 1M events above) are reconstructed
- 100k ZZ events; (BR for Z→ssbar) ~15%
- 15k ZZ events. After applying the Tight WP of the tagger:
- 5.4k events → 88/sqrt(5400) = 1.2σ

Back-of-the envelope estimates

THOROUGH STUDIES NEEDED

Summary & outlook

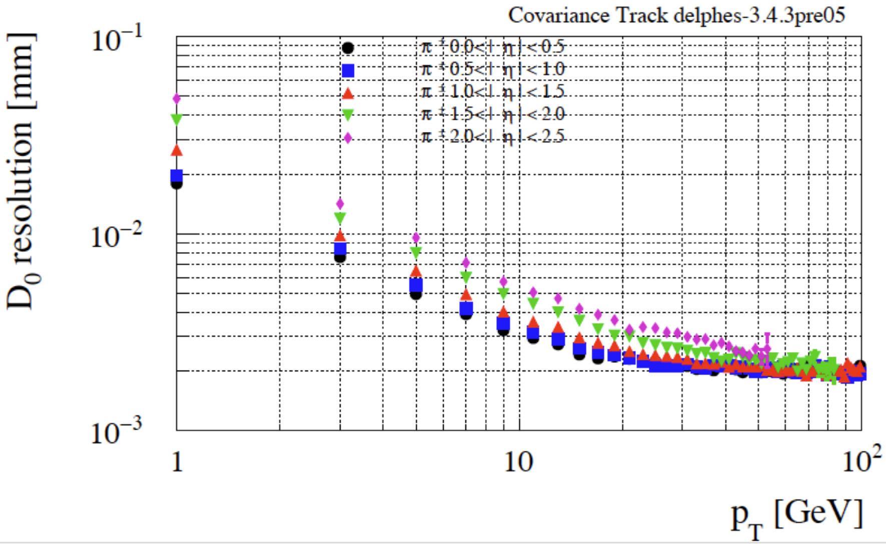
- Powerful jet flavour identification essential for the success of the e^+e^- physics program
- A first version of a jet identification algorithm based on **PF candidates** and **PID** and **advanced ML** in place
  - Multi-class classifier b/c/s/ud/g
    - Results promising, in particular for charm and strange tagging
- PRELIMINARY conclusions:
  - adding an additional vertex layer does not tremendously improve b-tagging performance (resolution of ~ 2um already outstanding)
    - but improves charm tagging
  - There seems to be room for improving strange tagging with more powerful PID

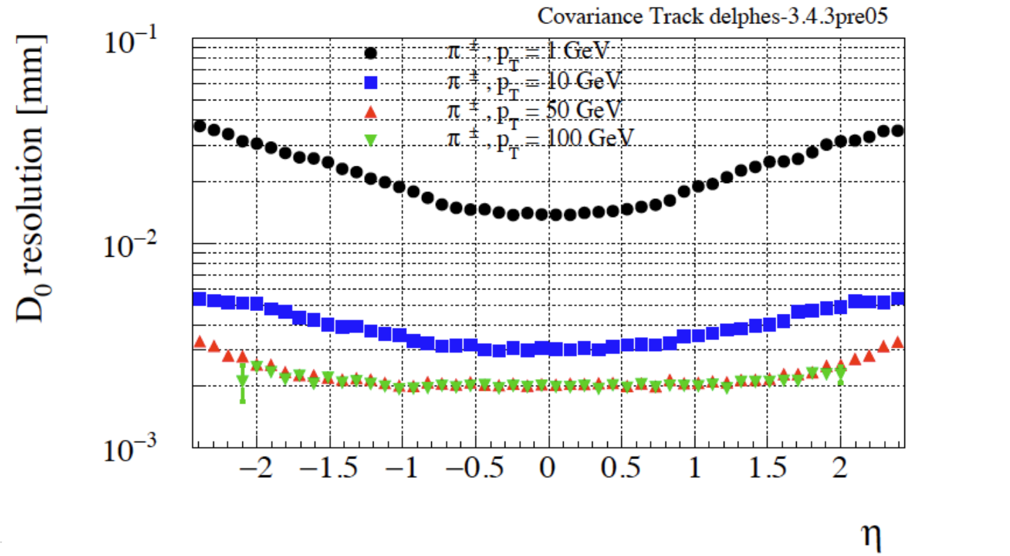
Backup

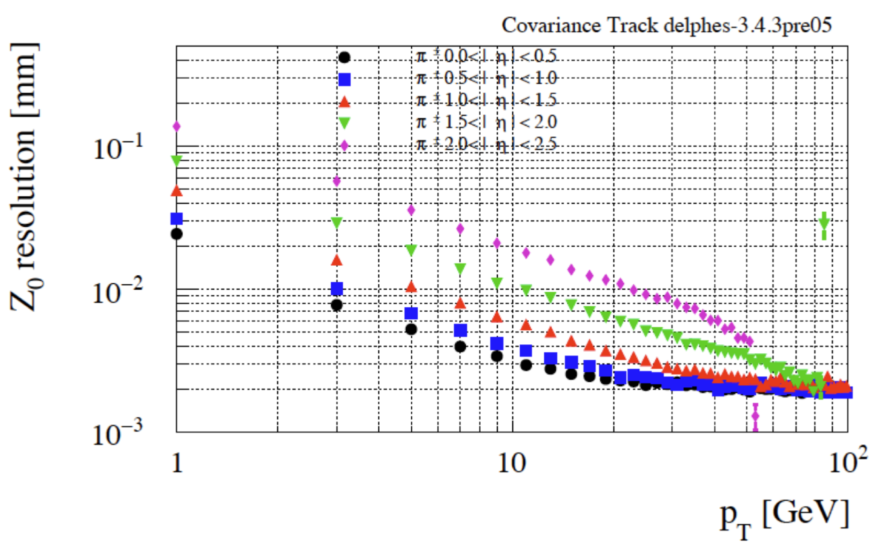
Impact parameter performance

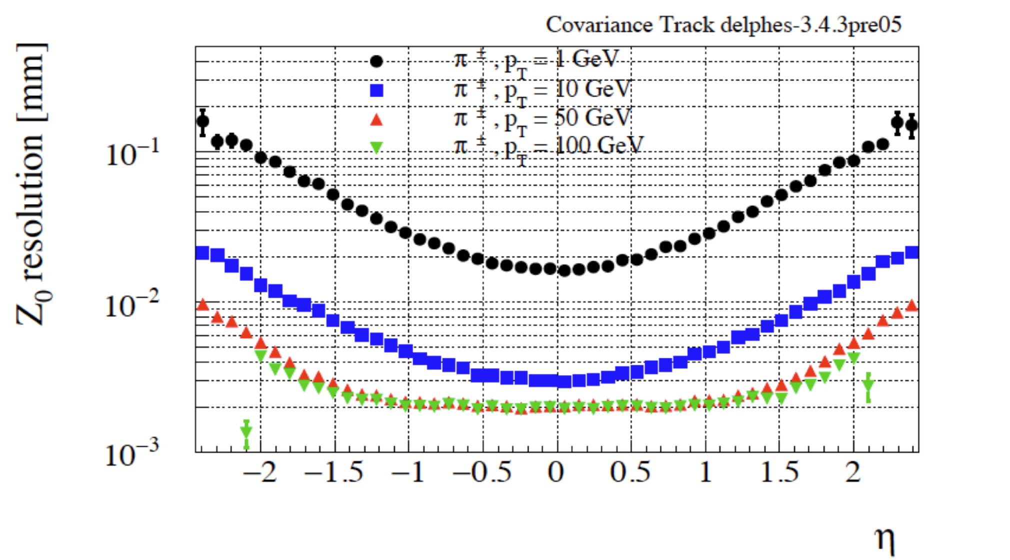
Credits to Sylvie Braibant

IDEA detector:







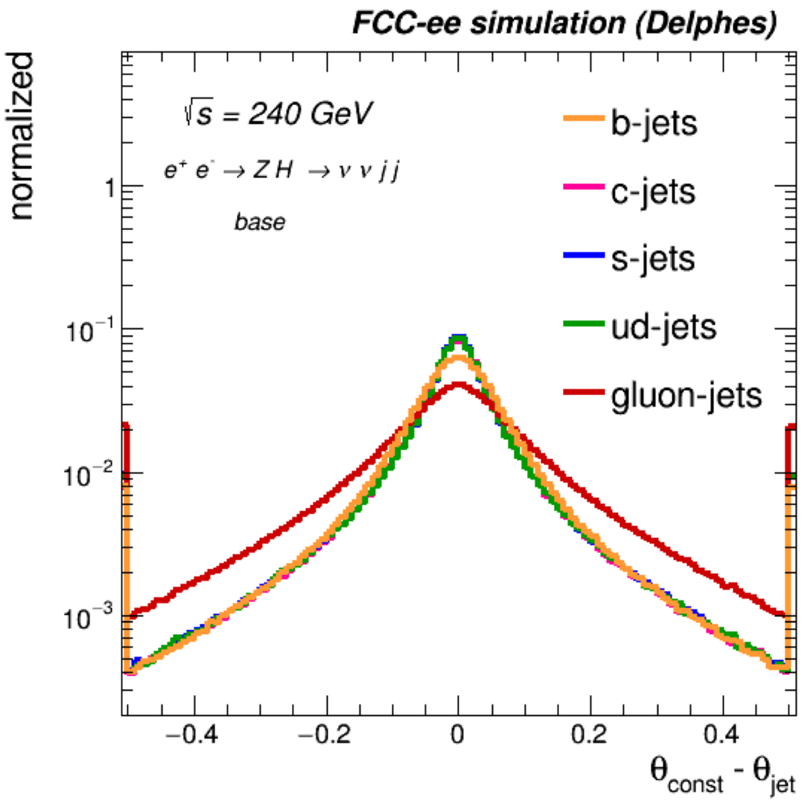


2μm IP resolution at high-p_T

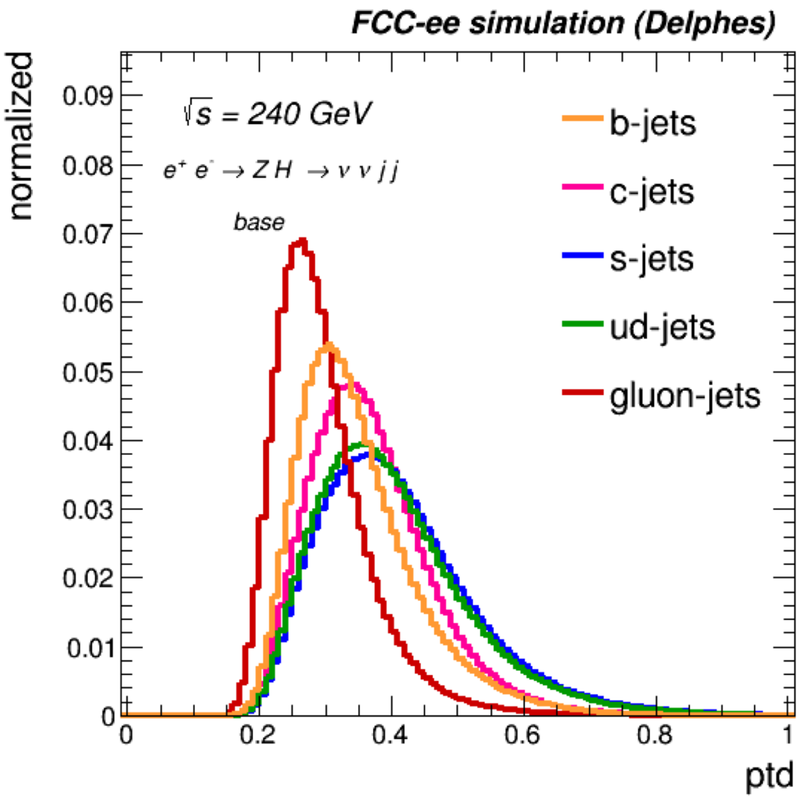
Input variables

- Comparison of input distributions for different jet flavors

Projection || to jet axis

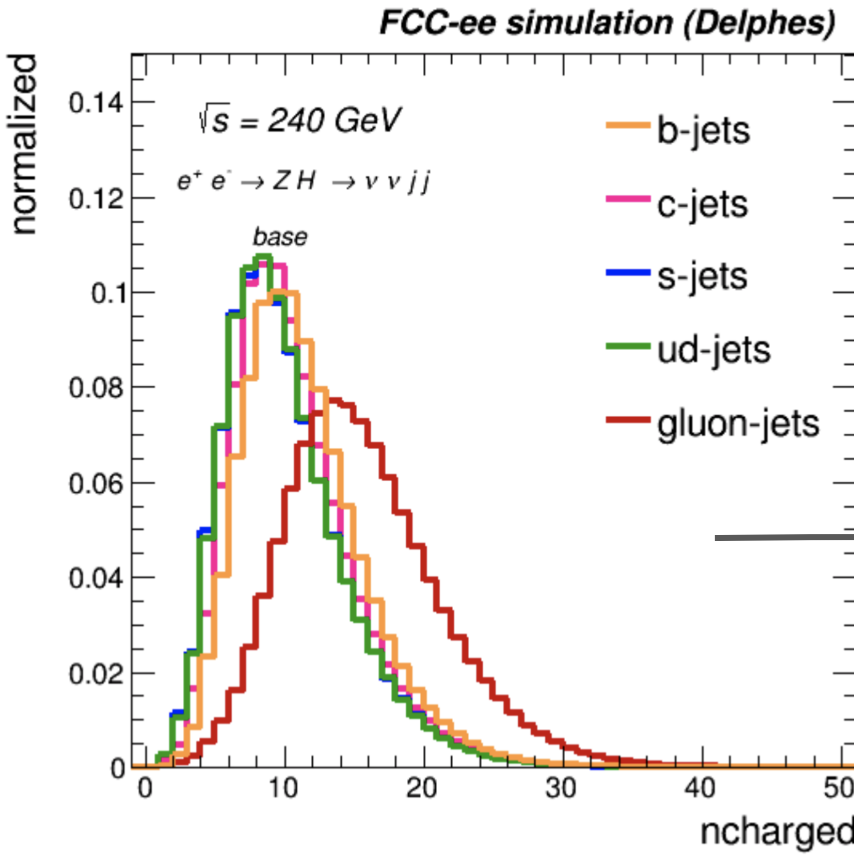


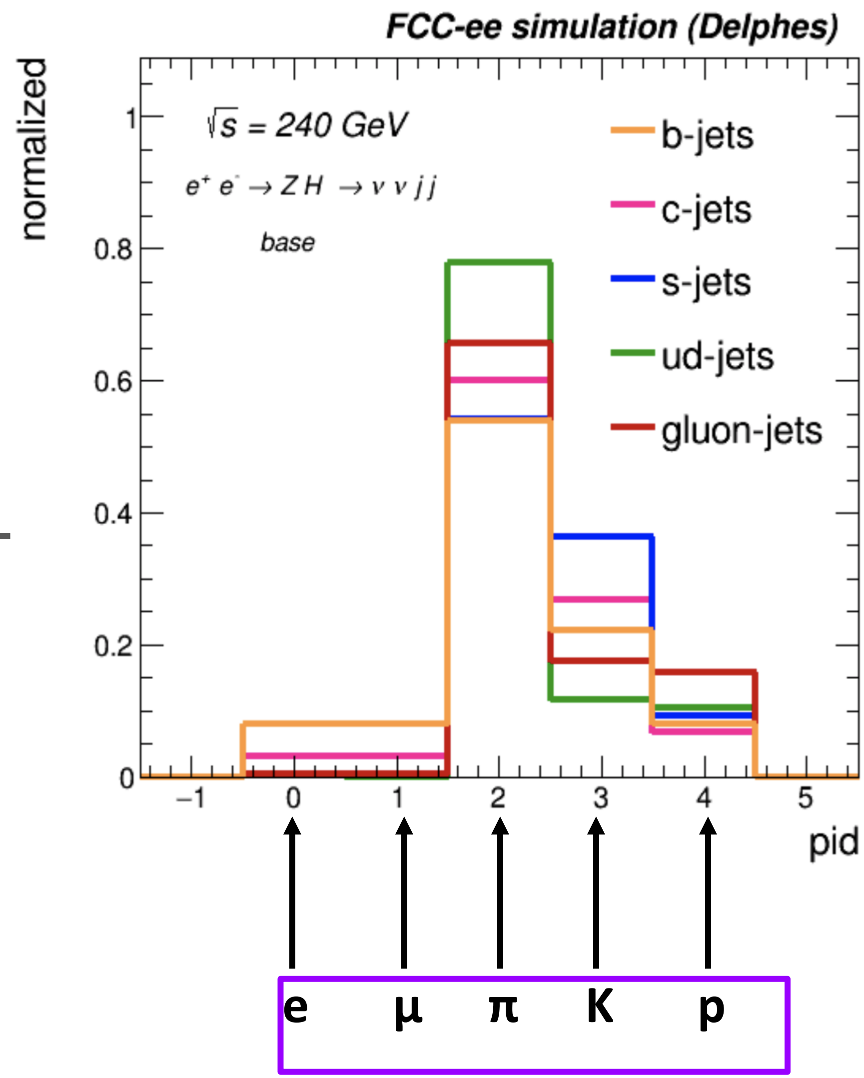
$p_T D$



- More comparisons: https://selvaggi.web.cern.ch/selvaggi/FCC/FCCee/FlavourTagging/

Performance w/ PID





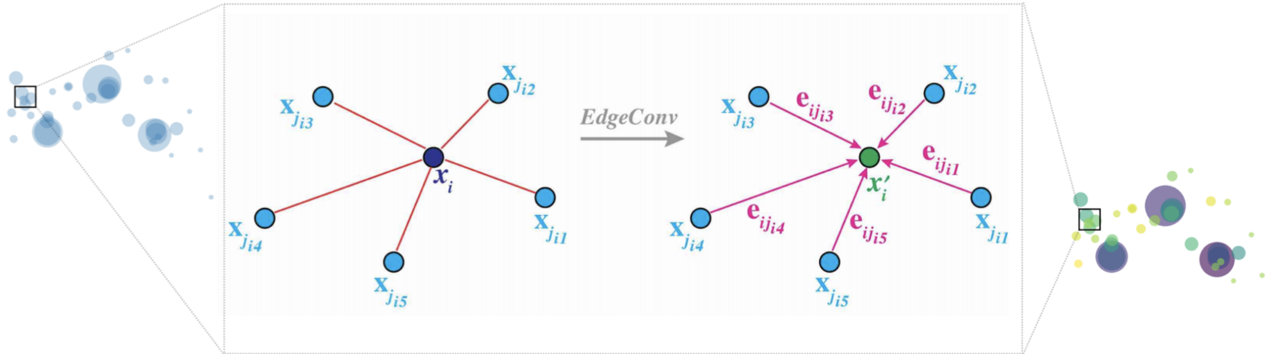
no PID: only charge

realistic: e, μ, mtof, dNdx

perfect PID: e, μ, + π, K, p from MC truth

Convolution on point cloud: EdgeConv

DGCNN: [arXiv:1801.07829]



EdgeConv: convolution on a graph

- **point cloud** is treated as **graph**, where each point is a **vertex**
- **local patch** defined by finding k-nearest neighbours
- **convolution** function:
  - define "edge feature" for each center-neighbour pair
    - $e_{ij} = h(x_i, x_j)$
  - aggregate all the features **symmetrically**:
    - $x'_i = \text{mean}_j \, e_{ij}$

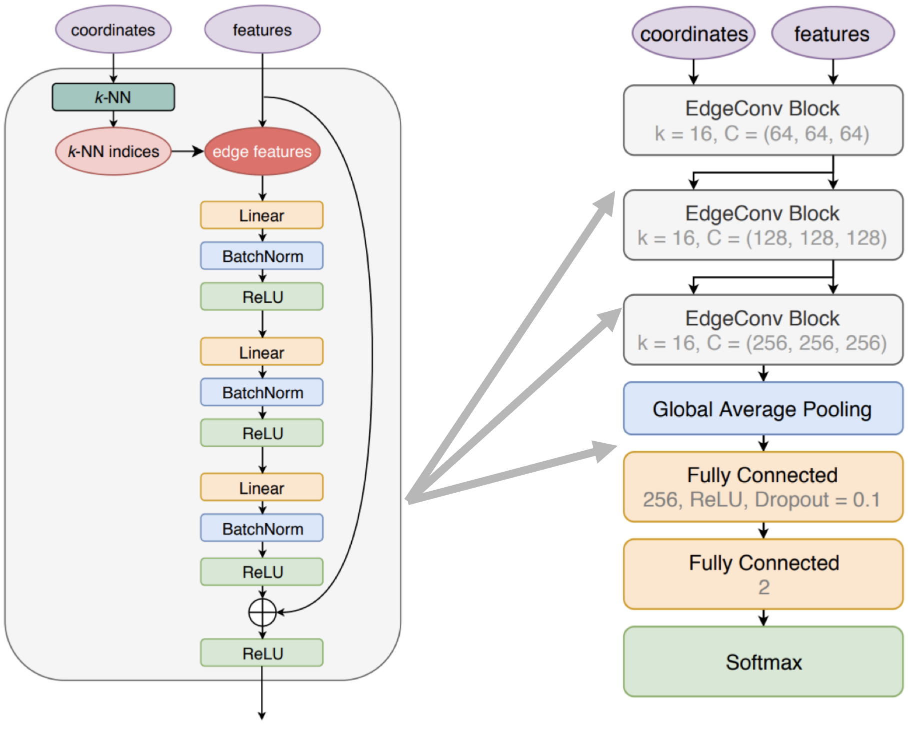
Key point:

Generalizing CNN for un-ordered/sparse images

ParticleNet

ParticleNet: [arXiv:1902.08570]

- **local neighborhood** information automatically incorporated
- **EdgeConv** layers can be **stacked** (as CNNs), and learn **local** (shallow layers) and **global** features (deep layers)
- **new features** provide new coordinates (in some abstract latent space) to compute “local patch” in new iteration



EdgeConv block   *ParticleNet architecture*

Designing a jet flavour tagging algorithm

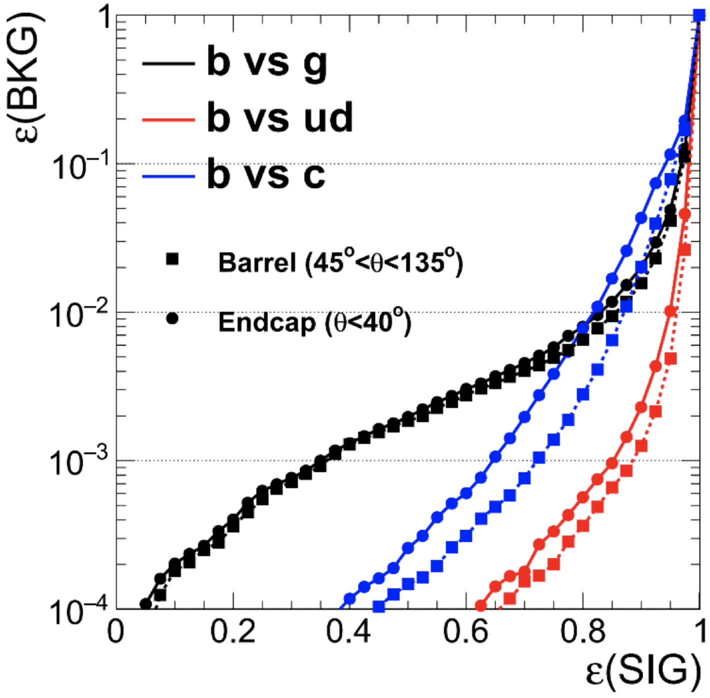
- How to represent a jet is one of the key aspects of algorithms for jet tagging
  - Improve performance → extend physics reach
  - Lead to fresh insight into jets → deepen our understanding of jet physics

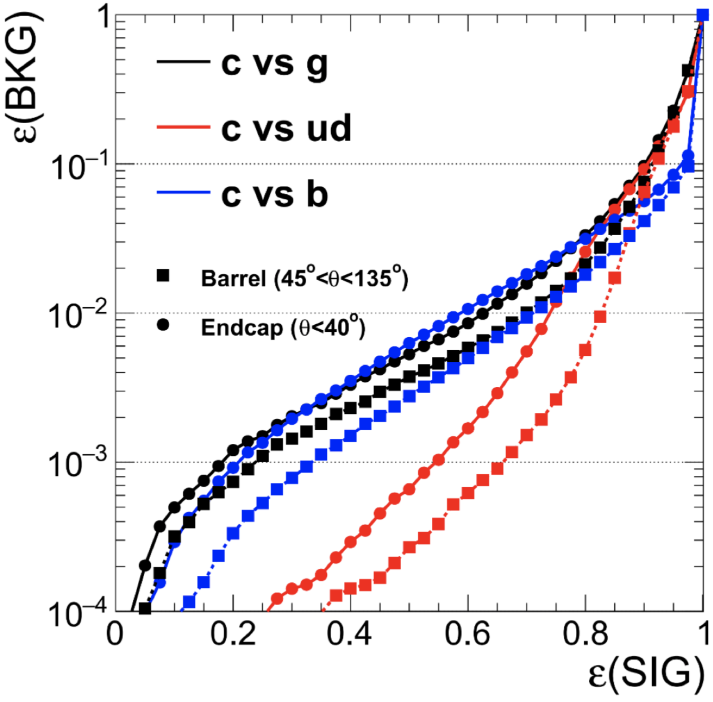
- Particles [associated to each jet] are intrinsically unordered
  - i.e., ordering by p_T(particle) or displacement from PV: suboptimal
  - Primary information: 2D coordinates in theta-phi space
  - Include additional features / particle: energy, displacement, charge, track quality, PID ...

Performance vs theta (b/c)

b-tagging | **c-tagging**

PRELIMINARY !! (LOW STATS TRAINING)





Tracking in Delphes

pattern recognition
track fitting

Full Simulation

-most accurate
-least flexible
-does not allow for change of geometry

parameterisation
$(d_0, d_z, p, \mathrm{ctg}\,\theta, \varphi)$

Delphes

TkLayout

-standalone tool
-predicts tracking performance for a given geometry
-allows for quick turnaroud
- no further implementation needed in Delphes
-requires intermediate step

FastTrackCovariance

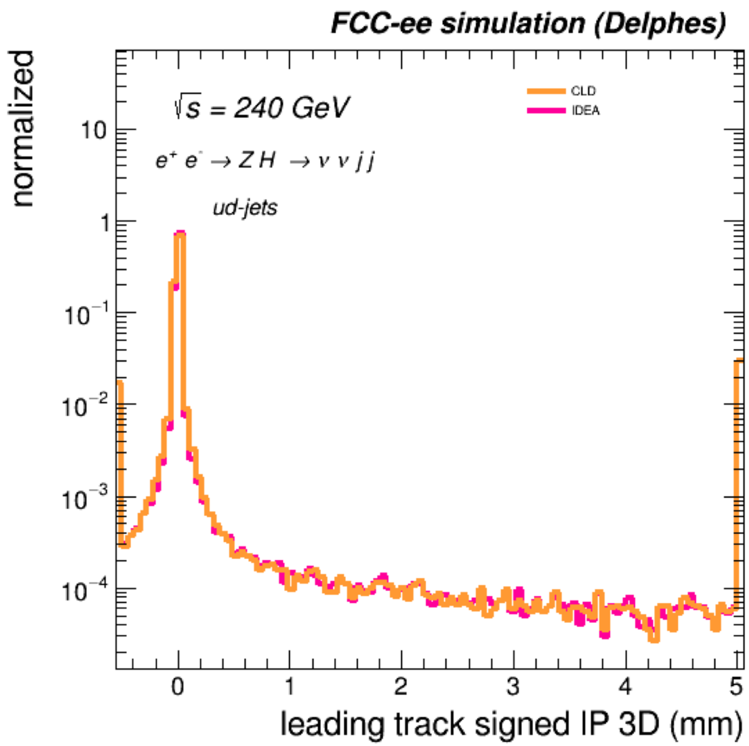
-(for now) standalone tool
-can be plugged into Delphes
-predicts performance given geometry
-allows for quick turnaroud
-not implemented in Delphes yet

23

Comparison: IDEA vs. CLD

- No big differences between in input variables between IDEA & CLD
  - small difference in material budget observed on light jets since dxy ~ 0
    - expect slightly better performance for IDEA detector for discrimination vs light

ud-jets



c-jets