Jet-flavour tagging at FCC-hh

FCC-hh Physics & Performance Meeting

Wei Sheng Lai, Nikita Pond, Tim Scanlon, Sebastien Rettie, Sam Van Stroud

17th October 2024

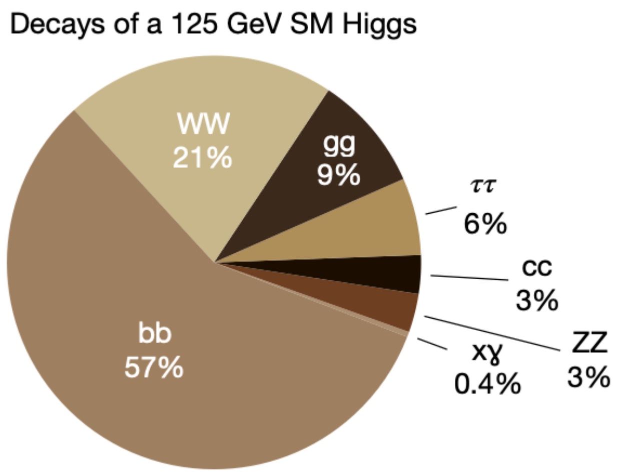
Overview

- Flavour tagging
  - Motivation
  - Transformer based model (first attempt @ FCC-hh)
- Validating workflow with FCC-ee
- Data simulation
  - Fast detector simulation (Delphes)
  - Input variables
- Performance
- Summary

Flavour tagging (b-tagging)

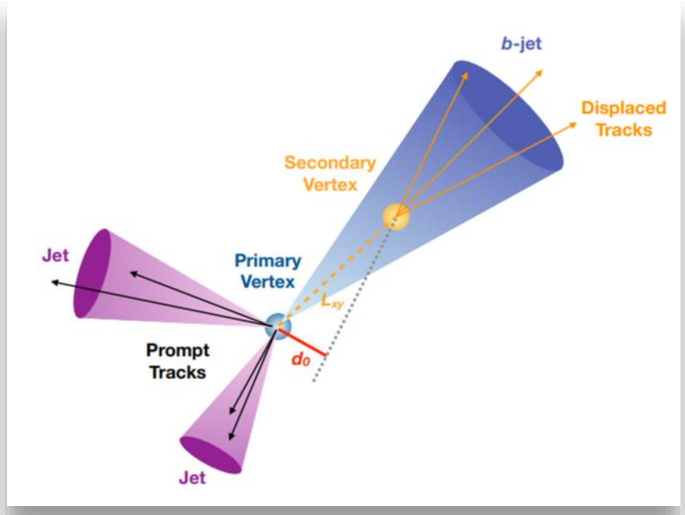
Why?

- Many signatures of interest contain b-jets
  - e.g Higgs boson decay
- Important in many new physics searches
- Also useful for rejecting backgrounds
  - e.g. $t\bar{t}$ production can be troublesome $(t \rightarrow bW)$



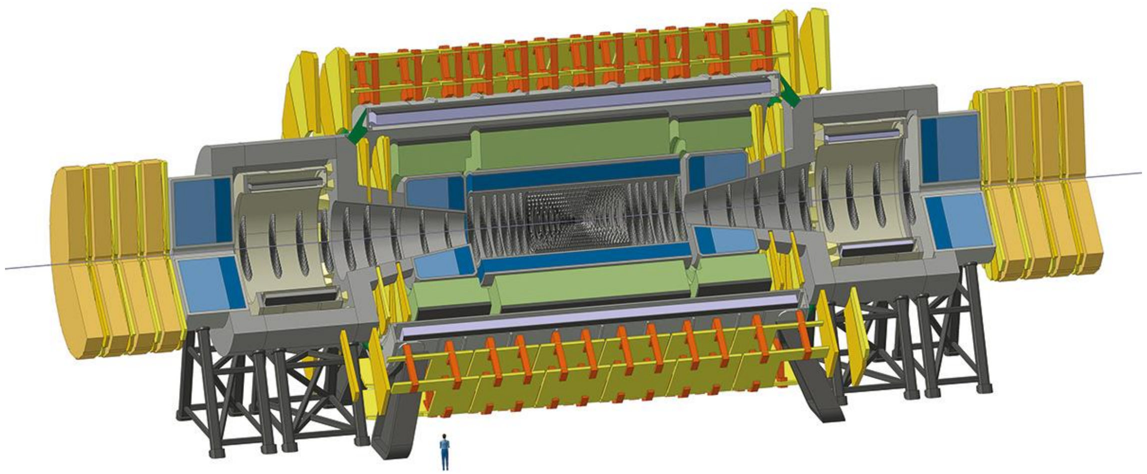
How?

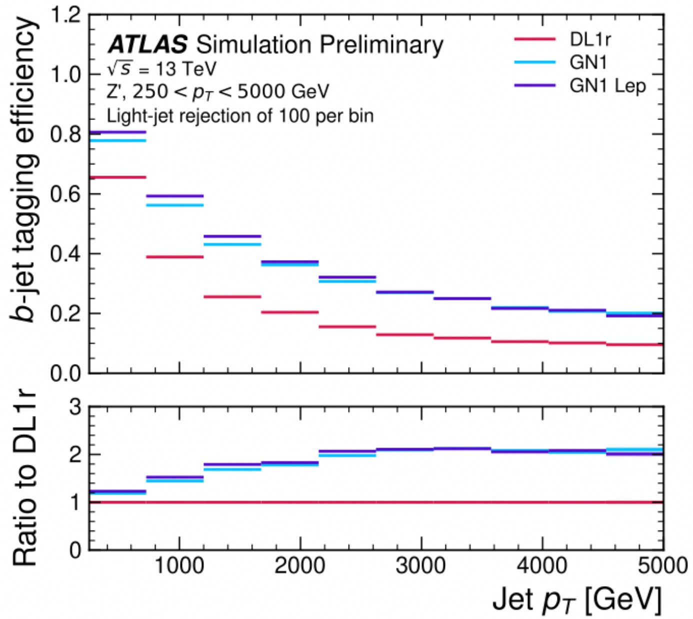
- b-hadrons have a relatively long lifetime (~1.5ps)
- Presence of 1+ displaced vertex within the jet
- Tracks with large impact parameter values (d_0)



Flavour tagging at FCC-hh

- Understand flavour-tagging performance at FCC-hh
  - Impact of new detector designs and collider conditions
- Study:
  - Performance as a function of pT and η
  - Impact on very large pile-up



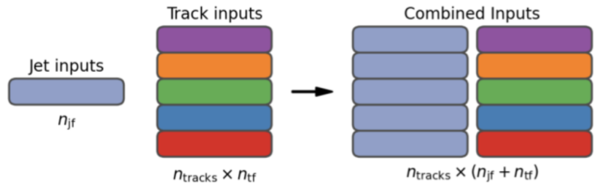


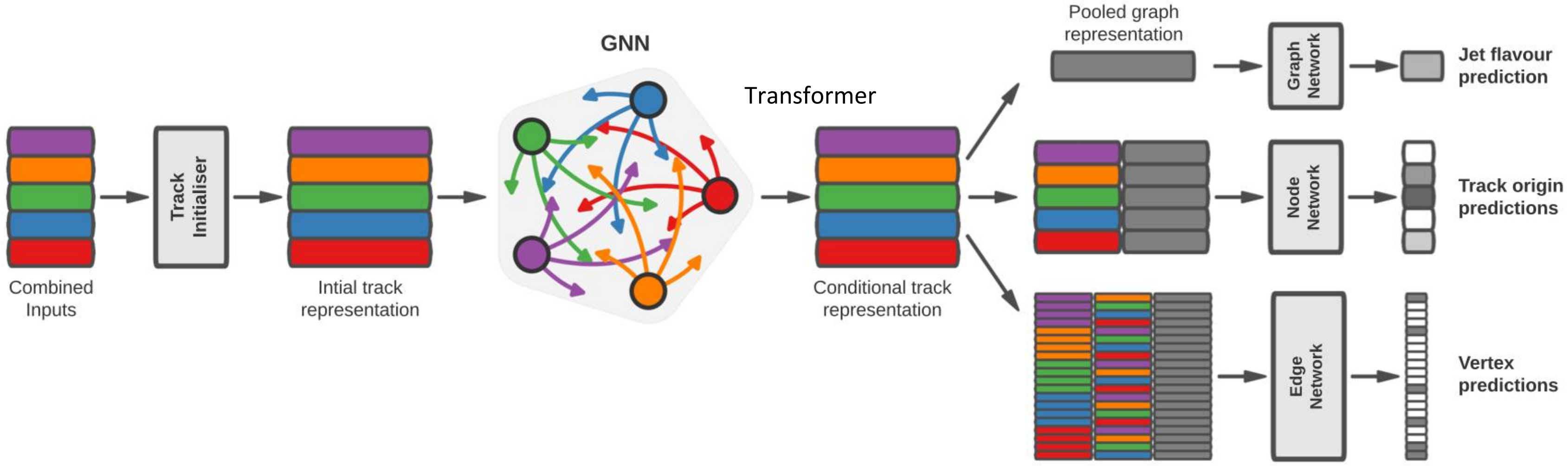
Ref

GN2

- Transformer-based architecture (GN2) [ATL-PHYS-PUB-2022-027]
- Jet kinematics are concatenated with jet constituents inputs
- Predict: Jet flavour classification (primary task),

  track truth origin and vertexing (auxiliary task)





Output:
p_i: probability of each classes

Workflow

UCL

Delphes + Pythia

- Event generation
- Detector simulation

DelphesFTAG

- ROOT to h5 ntuples (preferred format for ML training)
- Extracting auxiliary task information

Salt [link]

- Model training

Validate setup with FCC-ee

Inspired by Particle-Net [1902.08570] on FCC-ee [2202.03285]

- Authors (Michele Selvaggi) very helpfully provided their samples/code
- Use as benchmark to test setup, code generation and relative performance

Event generation (Madgraph + Pythia)

- $ee \rightarrow ZH \rightarrow \nu\nu jj$ ($\sqrt{s} = 240$ GeV)
- Provided by authors
- Focus on identifying jets from Higgs decays (b, c, s, ud, gluon)
- 10 million jets labelled according to the Higgs decay process

Performance

Discriminant Function constructed to study the discriminant between pairs of flavour

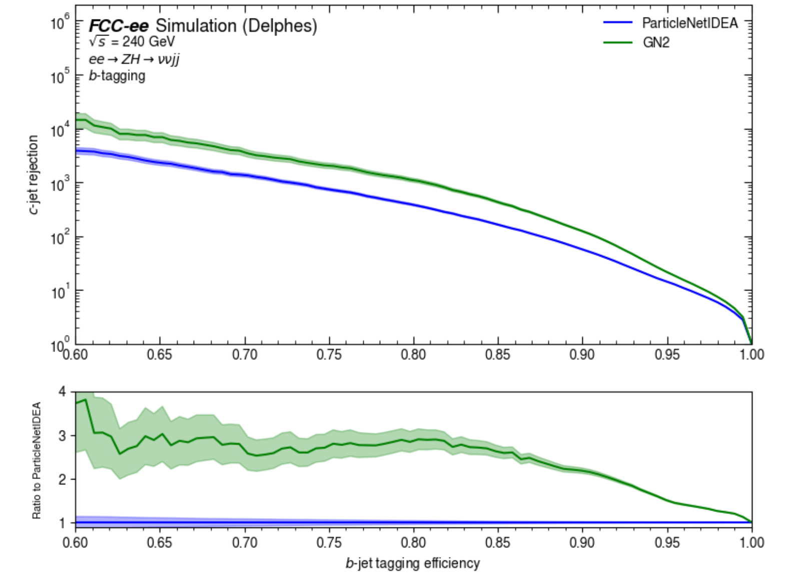
$$D_{ij} = \frac{p_i}{p_i + p_j}$$

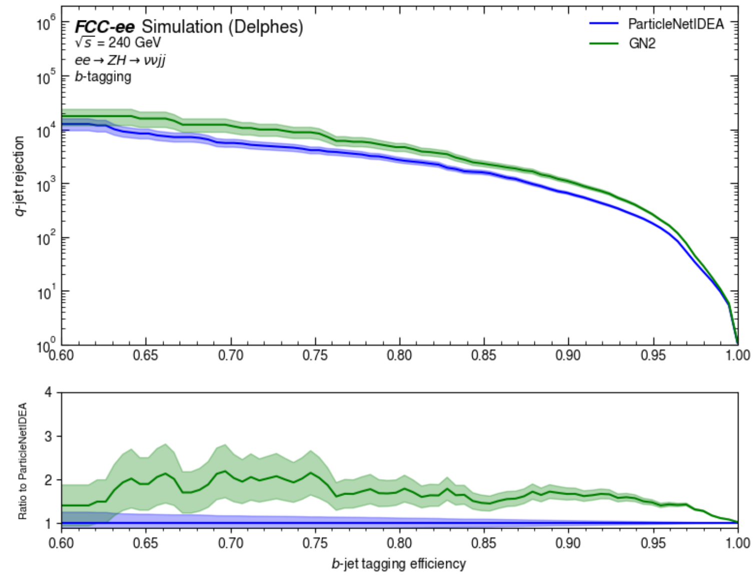
i = signal flavour
j = background flavour

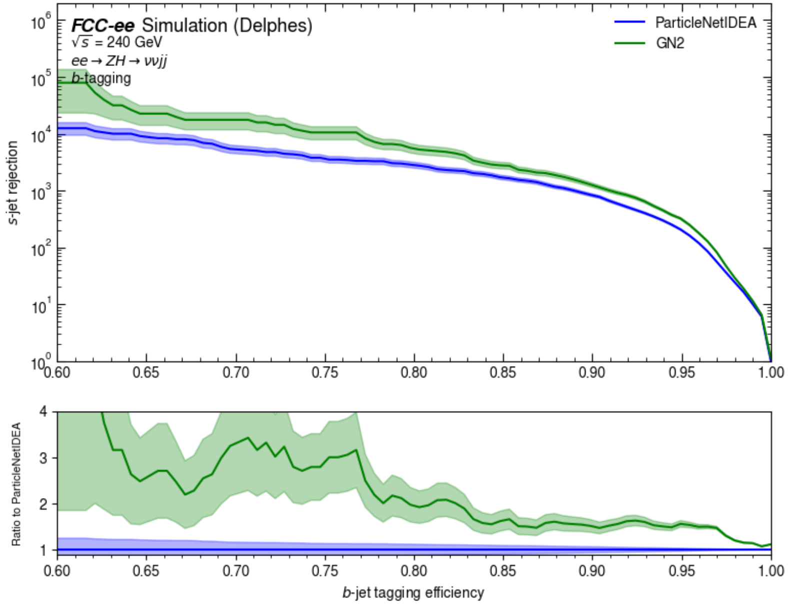
Use discriminant values cut to determined whether a jet is tagged or not

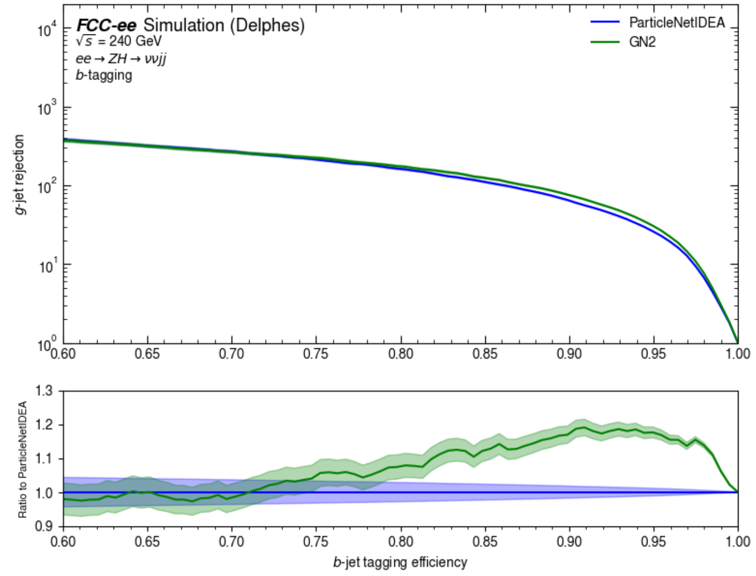
Results:

- Reproduced ParticleNetIDEA results
- Validated sample generation/evaluation
- Similar performance achieved



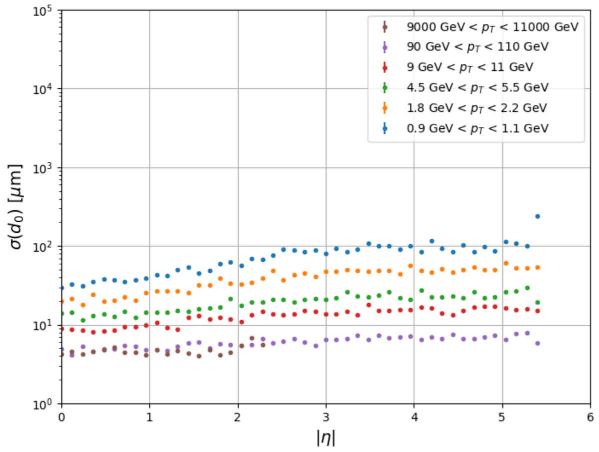




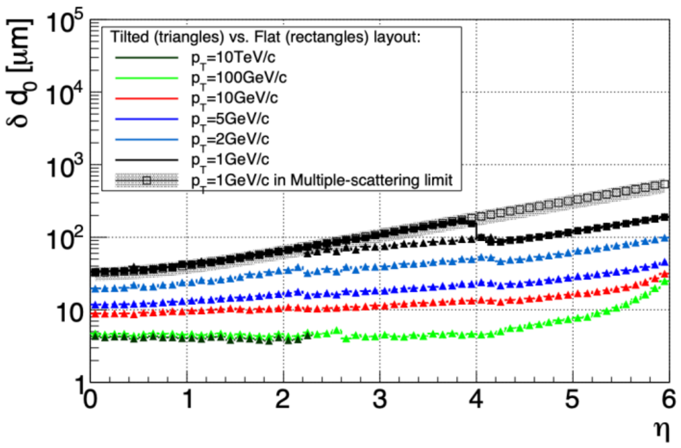


FCC-hh with Delphes

- Modified baseline FCC-hh Delphes cards based on conceptual design report (CDR) [CERN-2022-002]
- DenseTrackFilter module:
  - Remove overlapping tracks in dense environment
- TrackCovariance module:
  - Similar to FCC-ee IDEA detector approach
  - Implemented simple description of tracking detector layout (tklayout link)
  - Simulate track parameters and covariance matrix
  - Validated resolutions agree with predictions from CDR
- Same approach to Scenario I detector card (Optimistic: higher tracking efficiency)



FCChhTrackCov_II



CDR

Simulating data (Pythia)

- Low pT (20 – 300 GeV):
  - $pp \rightarrow ZH \rightarrow \nu\nu jj$ (4 million jets)
- High pT (300 – 5000 GeV)
  - $pp \rightarrow Z' \rightarrow jj$ (4 million jets)
- Only select leading 2 jets
- Assume 0 pileup effect
- Primarily interested in identifying b or c-jets while rejecting other flavours
- Labelling of jet flavour is done based on truth hadron content
  - Similar to ATLAS approach

Jet Flavour	Presence of b-Hadron	Presence of c-Hadron
b-jet	✓	-
c-jet	✗	✓
light-jet	✗	✗

Input variables

UCL

Jet Variables	Description
p_T	Jet transverse momentum
η	Jet pseudorapidity
Track Variables	**Description**
$E_{\text{const}}/E_{\text{jet}}$	Energy of jet constituent divided by the jet energy
θ_{eta}	Polar angle of the jet constituent, relative to the jet momentum
ϕ_{rel}	Azimuthal angle of the jet constituent, relative to the jet momentum
d_{xy}	Transverse impact parameter of the track
d_z	Longitudinal impact parameter of the track
$\text{SIP}_{2\text{D}}$	Signed 2D impact parameter of the track
$\text{SIP}_{2\text{D}}/\sigma_{2\text{D}}$	Signed 2D impact parameter significance of the track
$\text{SIP}_{3\text{D}}$	Signed 3D impact parameter of the track
$\text{SIP}_{3\text{D}}/\sigma_{3\text{D}}$	Signed 3D impact parameter significance of the track
$d_{3\text{D}}$	Jet track distance at their point of closest approach
$d_{3\text{D}}/\sigma_{d3\text{D}}$	Jet track distance significance at their point of closest approach
C_{ij}	Covariance matrix of track parameters
q	Electric charge of the particle
isMuon	If the particle is identified as muon
isElectron	If the particle is identified as electron
isPhoton	If the particle is identified as photon
isChargedHadron	If the particle is identified as a charged hadron
isNeutralHadron	If the particle is identified as a neutral hadron

Jet kinematics

Relative kinematics

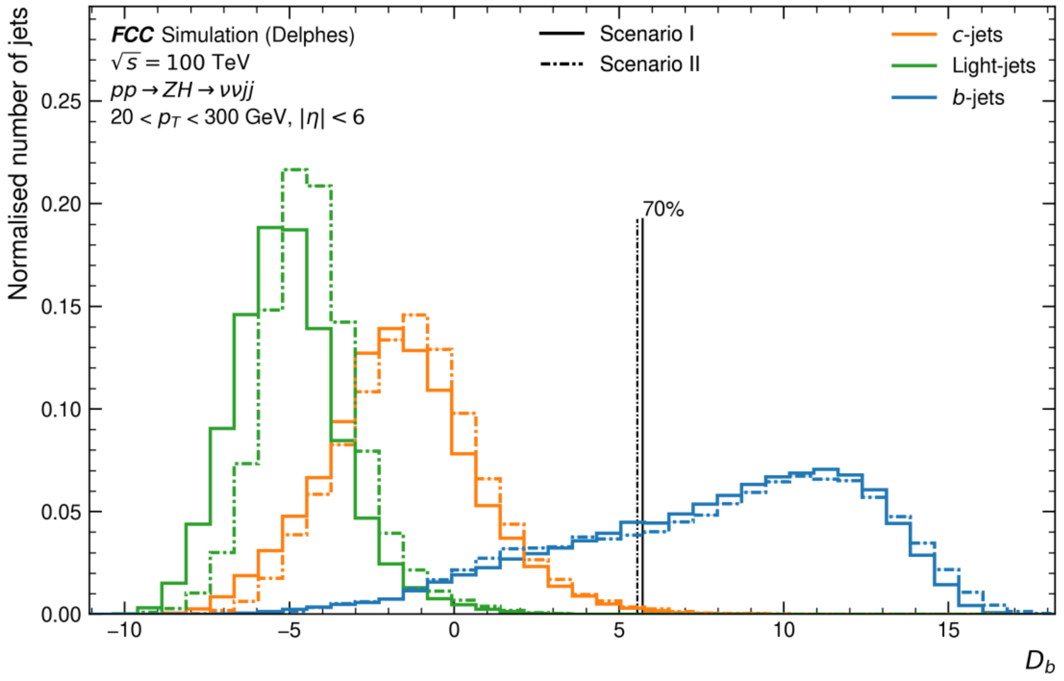
Track parameters

Particle Identification

ATLAS-style discriminant

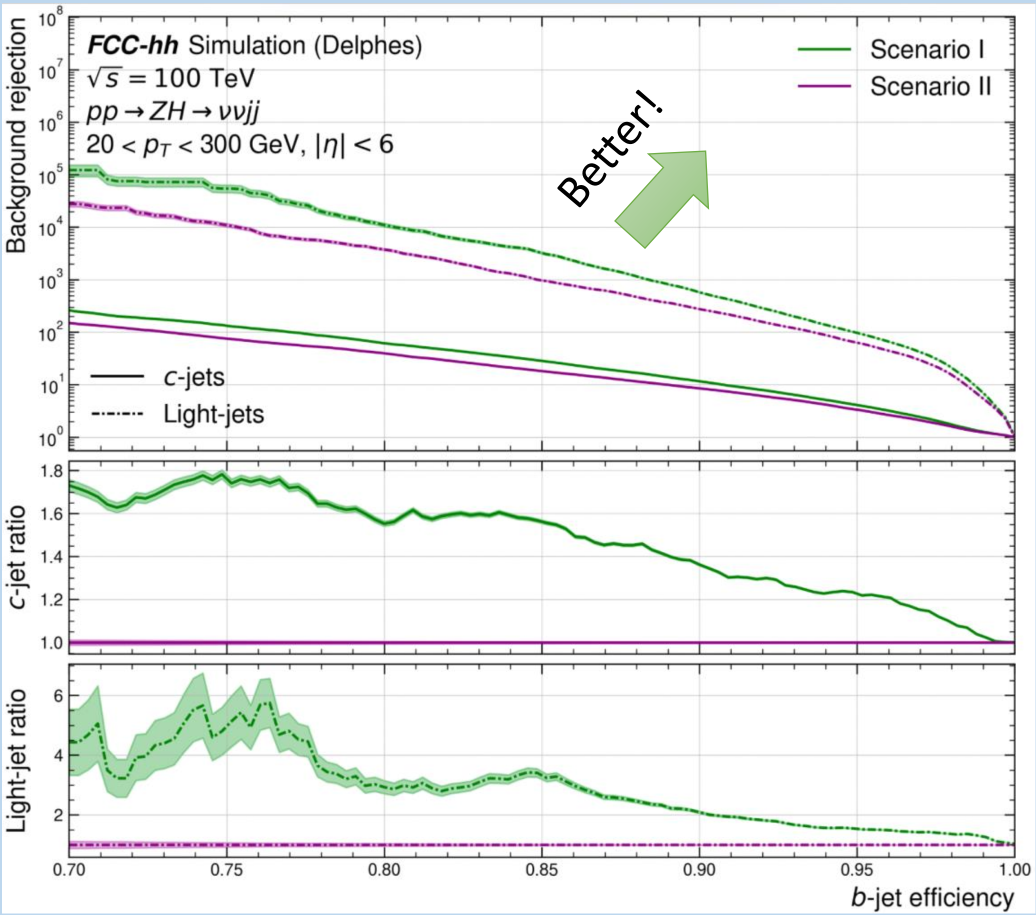
- Model outputs 3 probabilities p_b, p_c and p_l
- To identify b-jets while rejecting c and light-jets
  - Discriminant setup:

$$D_b = \ln\left(\frac{p_b}{f_c \cdot p_c + (1 - f_c) \cdot p_l}\right)$$

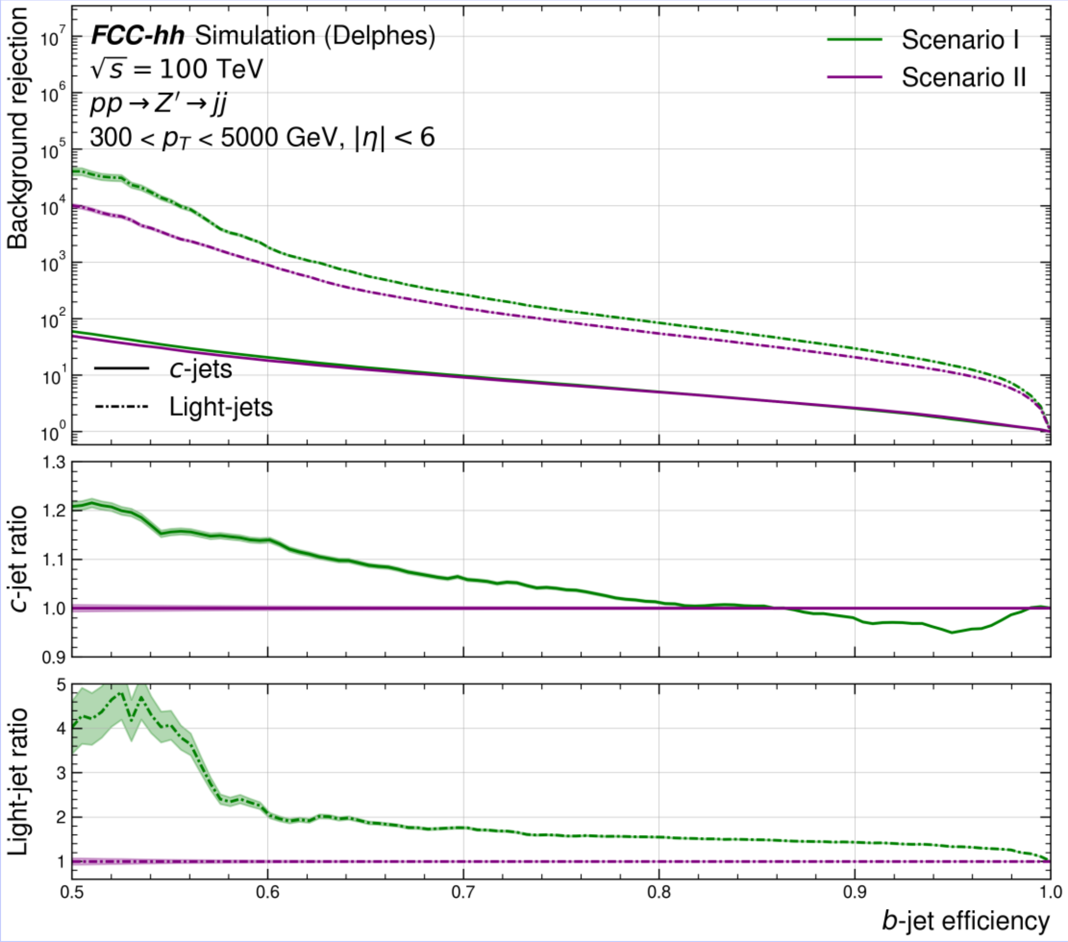


GN2 @ FCC-hh

Low pT (20 – 300 GeV)



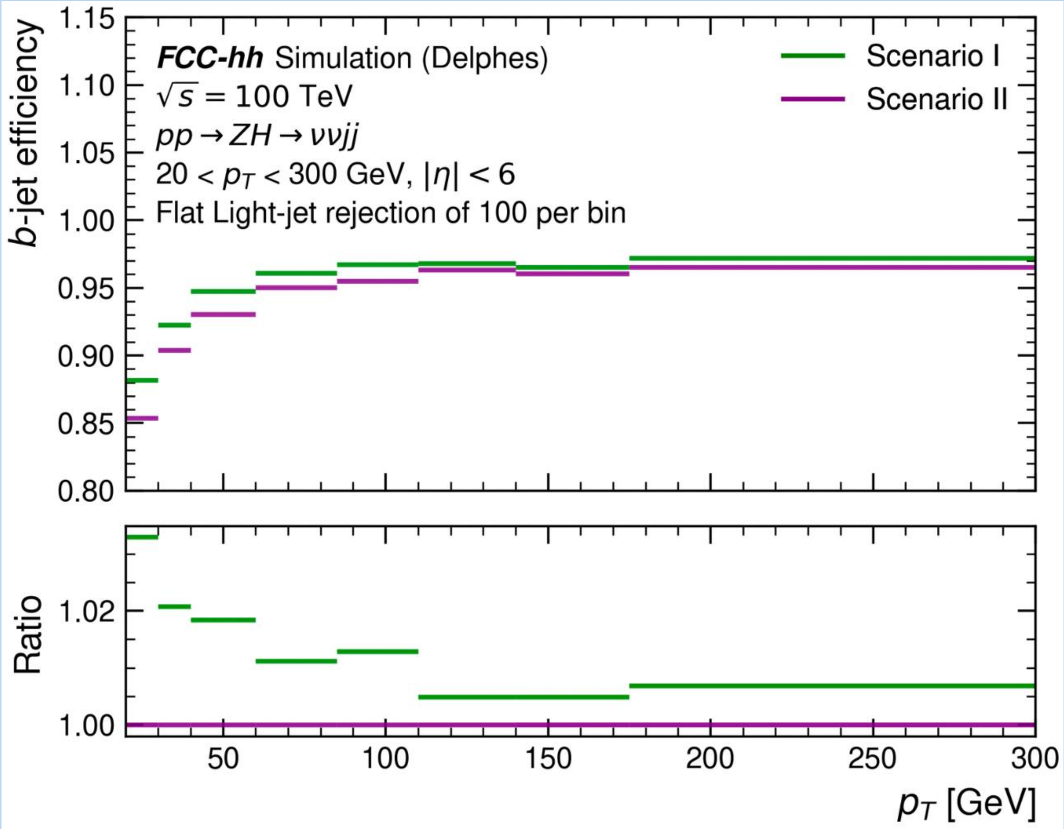
High pT (300 – 5000 GeV)



Very strong b-tagging performance!
With scenario I better than II (as expected)

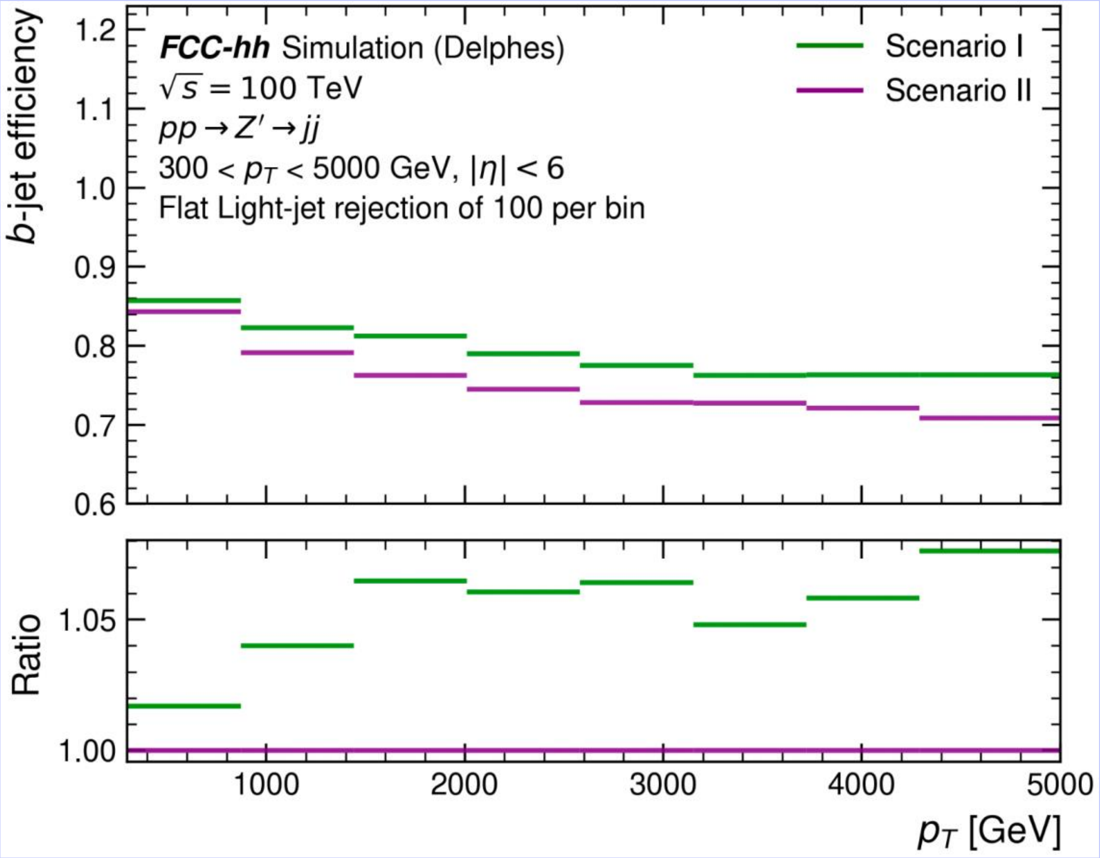
GN2 @ FCC-hh profiling performance

Low pT (20 – 300 GeV)



1% light jet mis-tagging rate
Reaches b-tagging efficiencies > 95% for moderate pT range

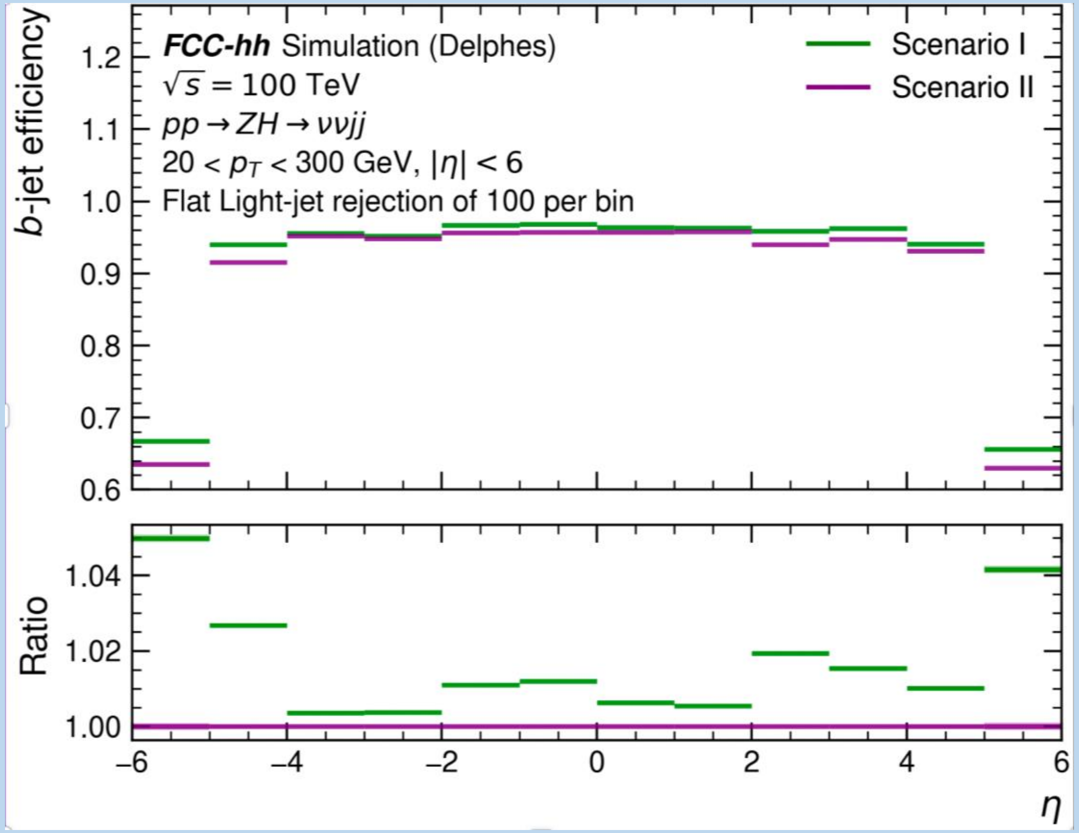
High pT (300 – 5000 GeV)



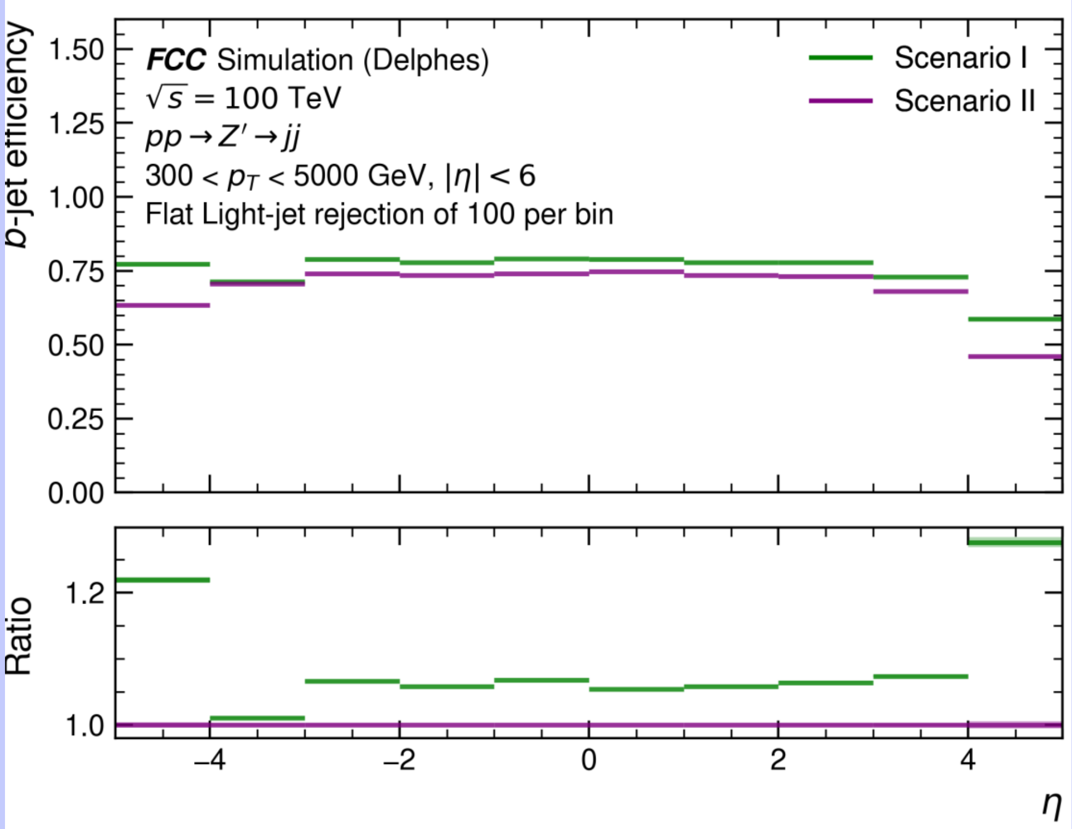
1% light jet mis-tagging rate
> 70% b-tagging efficiency for jets in TeV pT range

GN2 @ FCC-hh profiling performance

Low pT (20 – 300 GeV)



High pT (300 – 5000 GeV)



Flavour tagging performance maintained up to $|\eta| < 5$

Summary

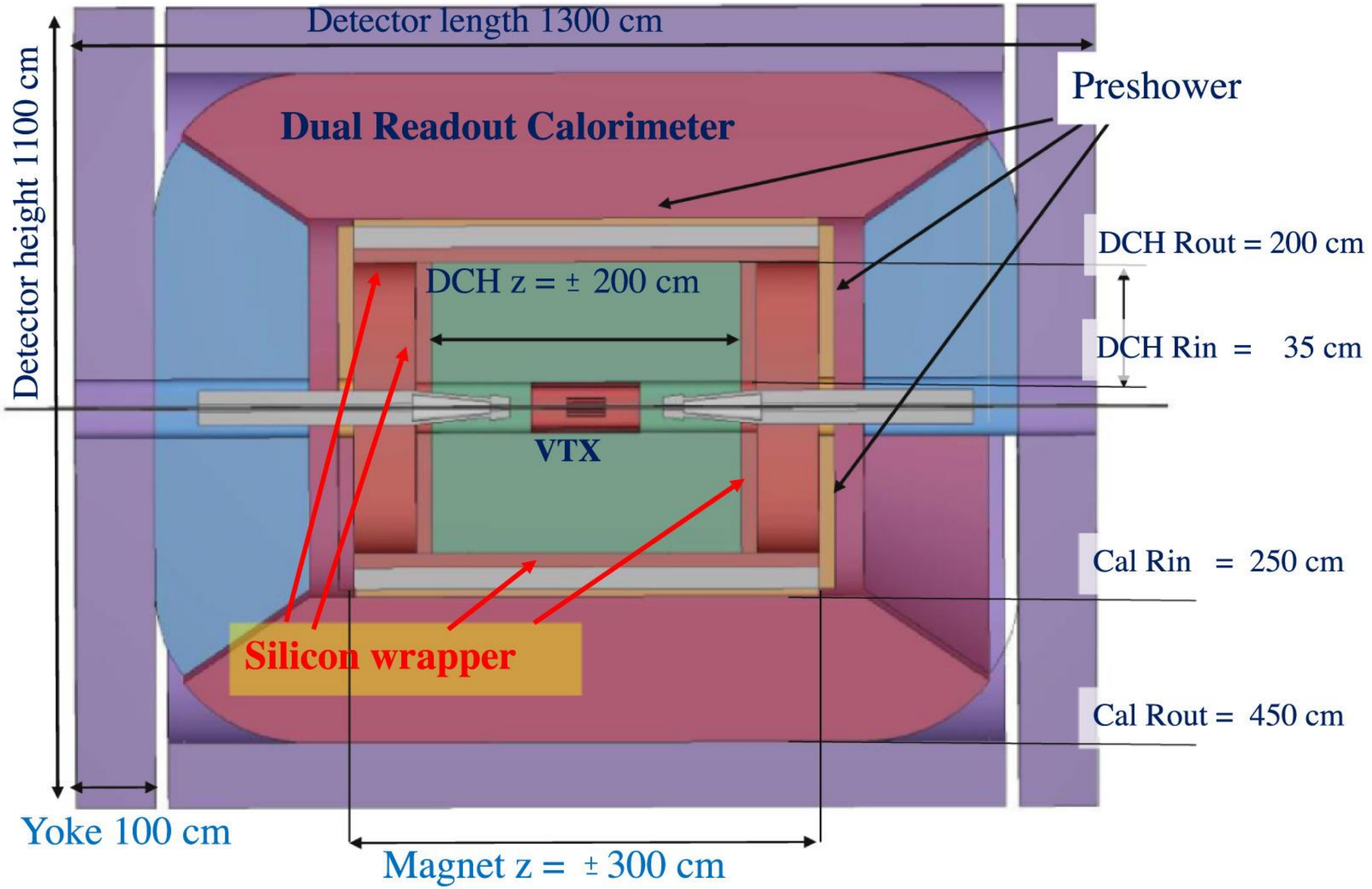
- Workflow established to study flavour-tagging performance at FCC-hh
- Studied performance with different detector scenarios (I/II)
- Provided parameterisations of b-tagging performance:

1. 1% light jet mis-tagging rate
   Obtain b-tagging efficiency > 95% for moderate pT jets
2. 1% light jet mis-tagging rate
   Maintain b-tagging efficiency > 70% for jets in TeV pT range
3. Maintain performance up to $|\eta| < 5$

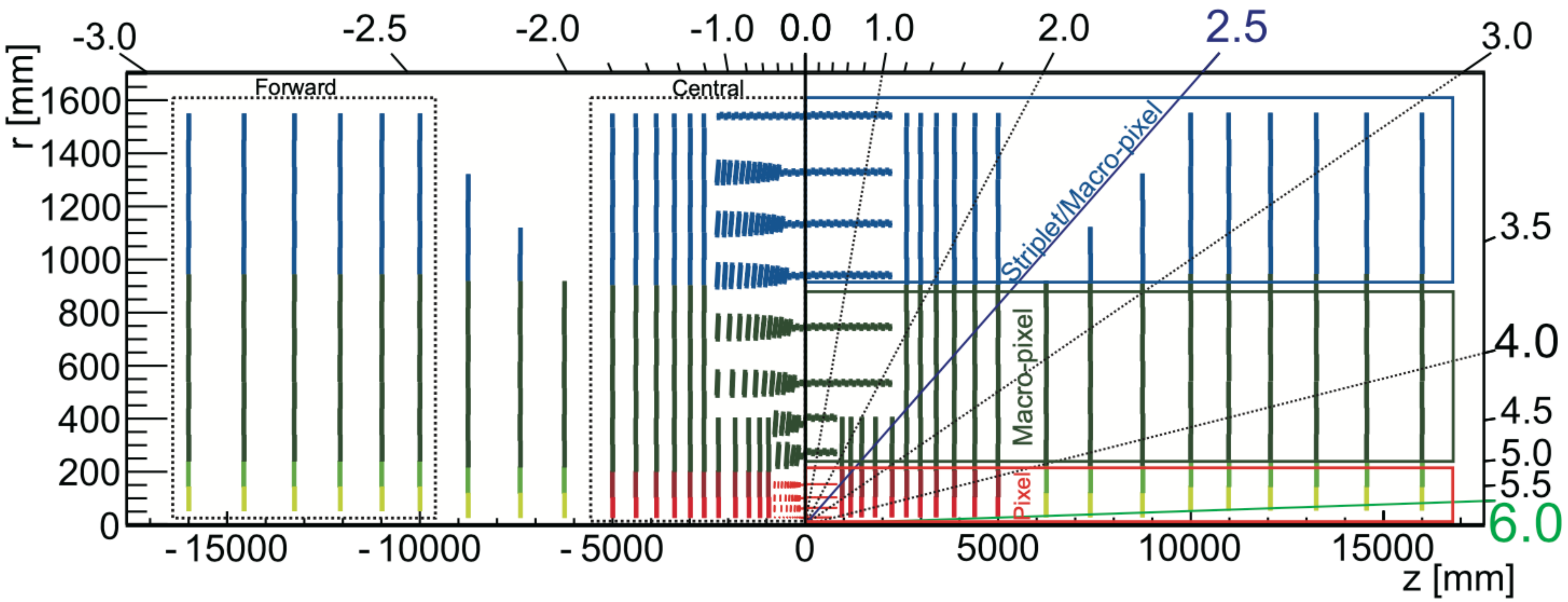
Interesting study: Adding pile-up

Backup

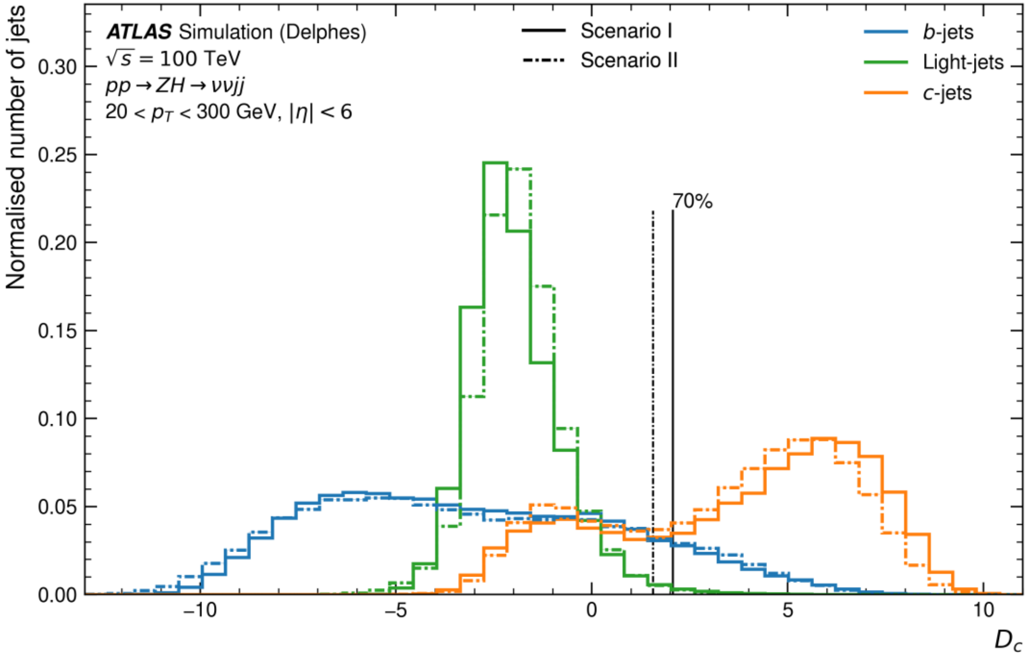
FCC-ee IDEA detector

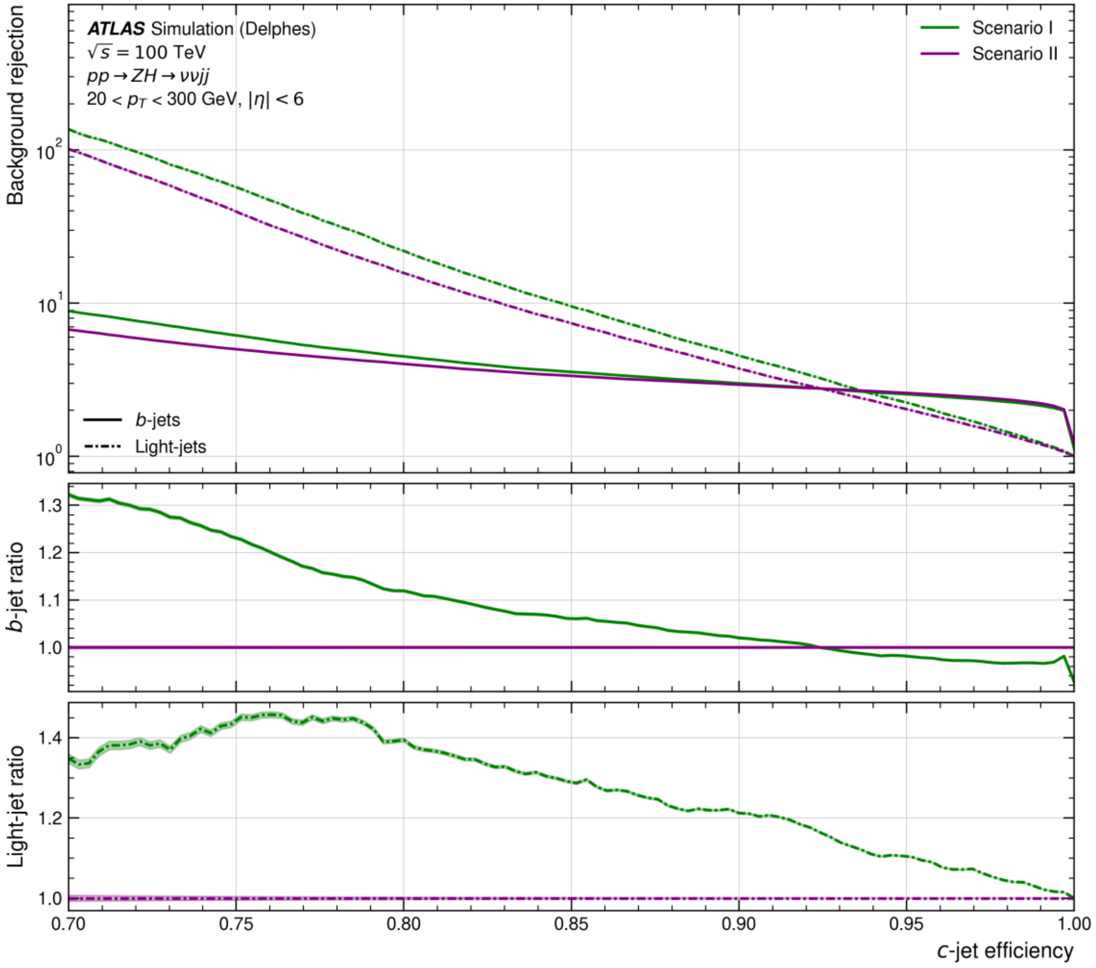


FCC-hh tracking detector



GN2 @ FCC-hh (c-tagging)

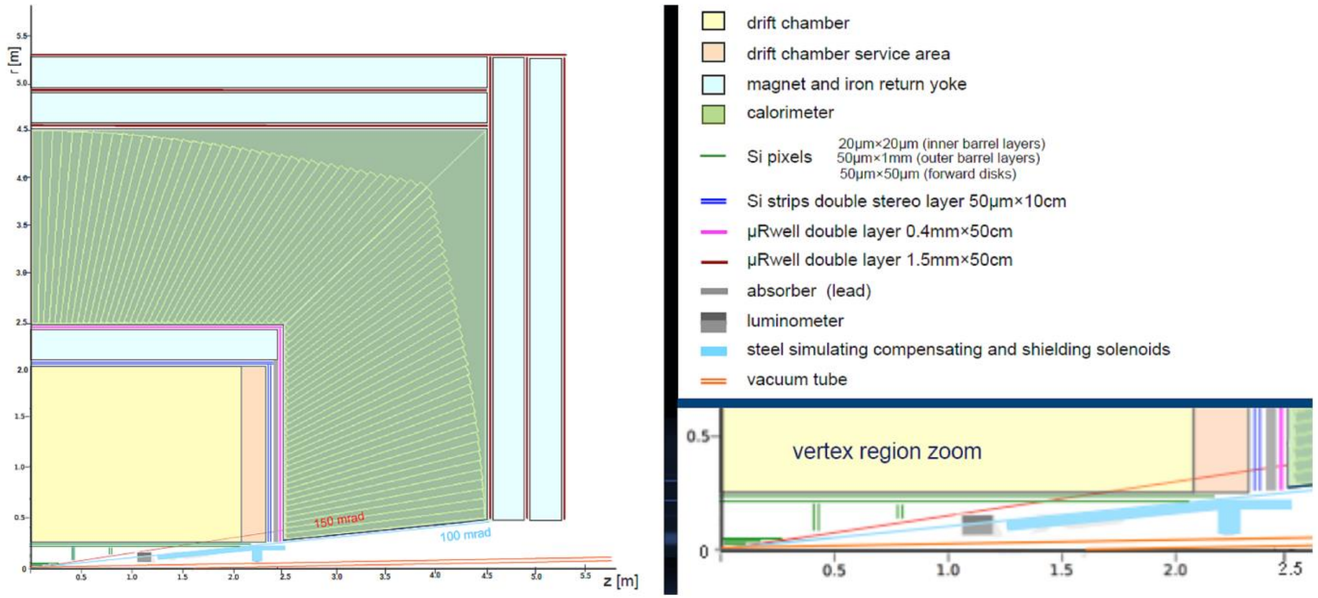




IDEA vs FCC-hh 1st layer

IDEA detector silicon pixel pitch size:

- 20um x 20um
- Resolution of 3um x 3um



FCChh detector silicon pixel pitch size:

- 25um x 50um
- Resolution of 7um x 14 um

Flat layout:

Pixels (inner)	Macro-pixels (middle)	Striplets/Macro-pixels (outer)
$25 \times 50\,\mu m^2$ (1–4th BRL) $25 \times 50\,\mu m^2$ (1st EC ring) $33.3 \times 100\,\mu m^2$ (2nd EC ring) $33.3 \times 400\,\mu m^2$ (3–4th EC ring)	$33.3 \times 400\,\mu m^2$	$33.3\,\mu m \times 50\,mm$ (BRL) $33.3\,\mu m \times 10\,mm$ (EC)

IDEA vs FCC-hh 1st layer

Inner layer of IDEA assumed to be closer to collision point

TABLE I. – *The main parameters of the IDEA concept detector.*

Parameters	
vertex technology	silicon
vertex inner/outer radius (cm)	1.7/34
tracker technology	drift chamber and silicon wrapper
tracker half length (m)	2.0
tracker outer radius (m)	2.0
solenoid field (T)	2.0
solenoid bore radius/half length (m)	2.1/3.0
preshower absorber	lead
preshower R_{min}/R_{max} (m)	2.4/2.5
DR calorimeter absorber	copper
DR calorimeter R_{min}/R_{max} (m)	2.5/4.5
overall height/length (m)	11/13

IDEA detector

Layer no :	1	2	3	4	5	6	Total
Average radius [mm] :	25.00	60.00	100.00	150.00	270.00	400.00	
Radius-min [mm] :	23.28	58.28	98.28	148.28	261.07	391.07	
Radius-max [mm] :	27.47	63.03	102.52	152.26	280.11	409.73	
Z-min [mm] :	-685.0	-820.0	-820.0	-820.0	-820.0	-820.0	
Z-max [mm] :	685.0	820.0	820.0	820.0	820.0	820.0	
Number of rods :	14	16	26	38	34	50	
Number of modules per rod :	20	40	40	40	17	17	
Number of modules :	280	640	1040	1520	578	850	**4908**

Disk no :	1	2	3	4	5	Total (+Z & -Z)
Radius-min [mm] :	25.0	25.0	25.0	25.0	25.0	
Radius-max [mm] :	404.0	404.0	404.0	404.0	404.0	
Average Z pos. [mm] :	950.0	1178.5	1462.0	1813.7	2250.0	
Z-min [mm] :	941.1	1169.6	1453.1	1804.8	2241.1	
Z-max [mm] :	958.9	1187.5	1471.0	1822.6	2258.9	
Number of rings :	4	4	4	4	4	
Number of modules per disk :	108	108	108	108	108	**1080**

Ring no :	1	2	3	4
R-min [mm] :	25.0	101.3	198.9	302.6
R-max [mm] :	104.7	204.6	303.4	404.0
Number of modules per ring :	12	20	32	44

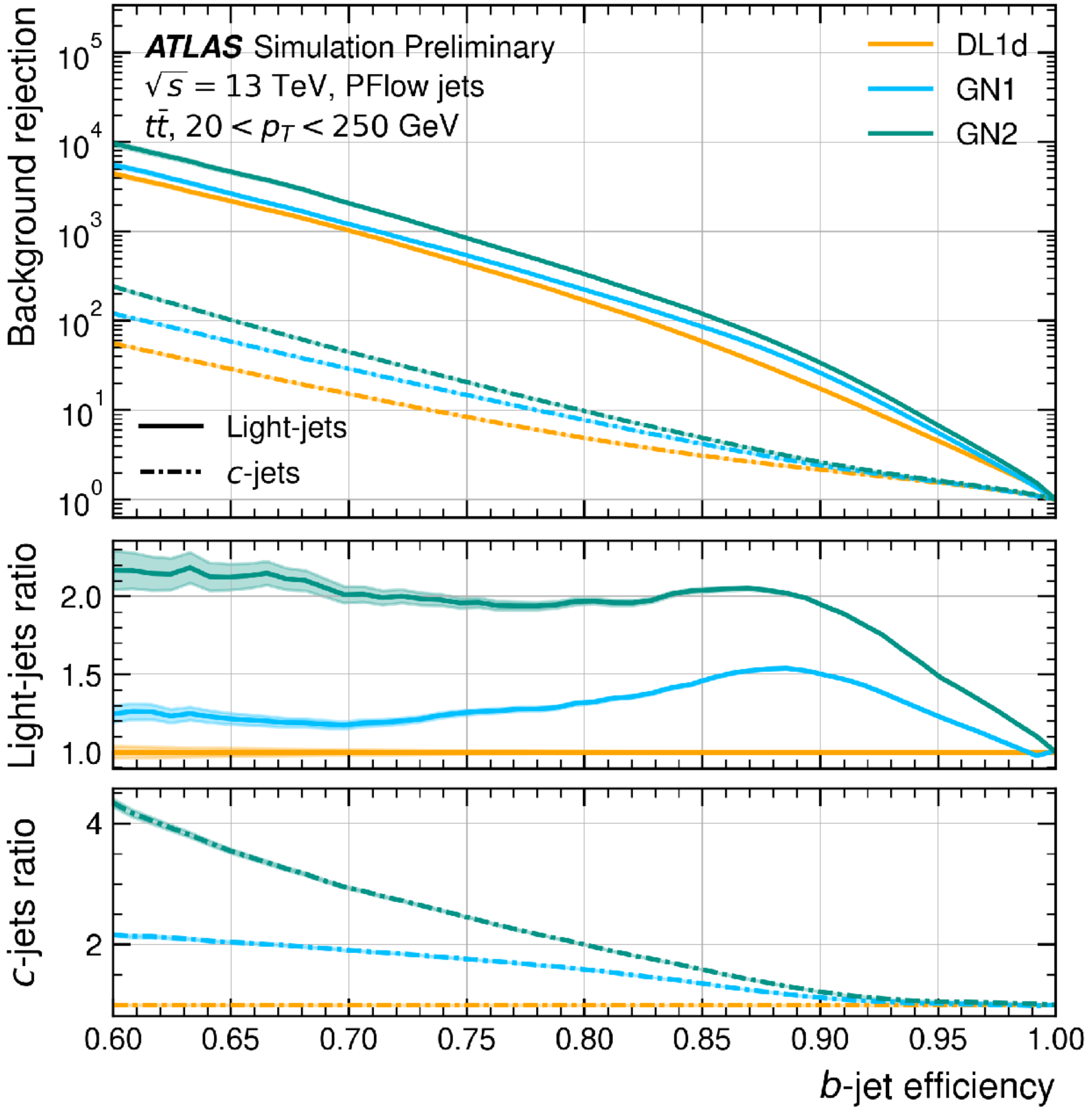
FCC detector

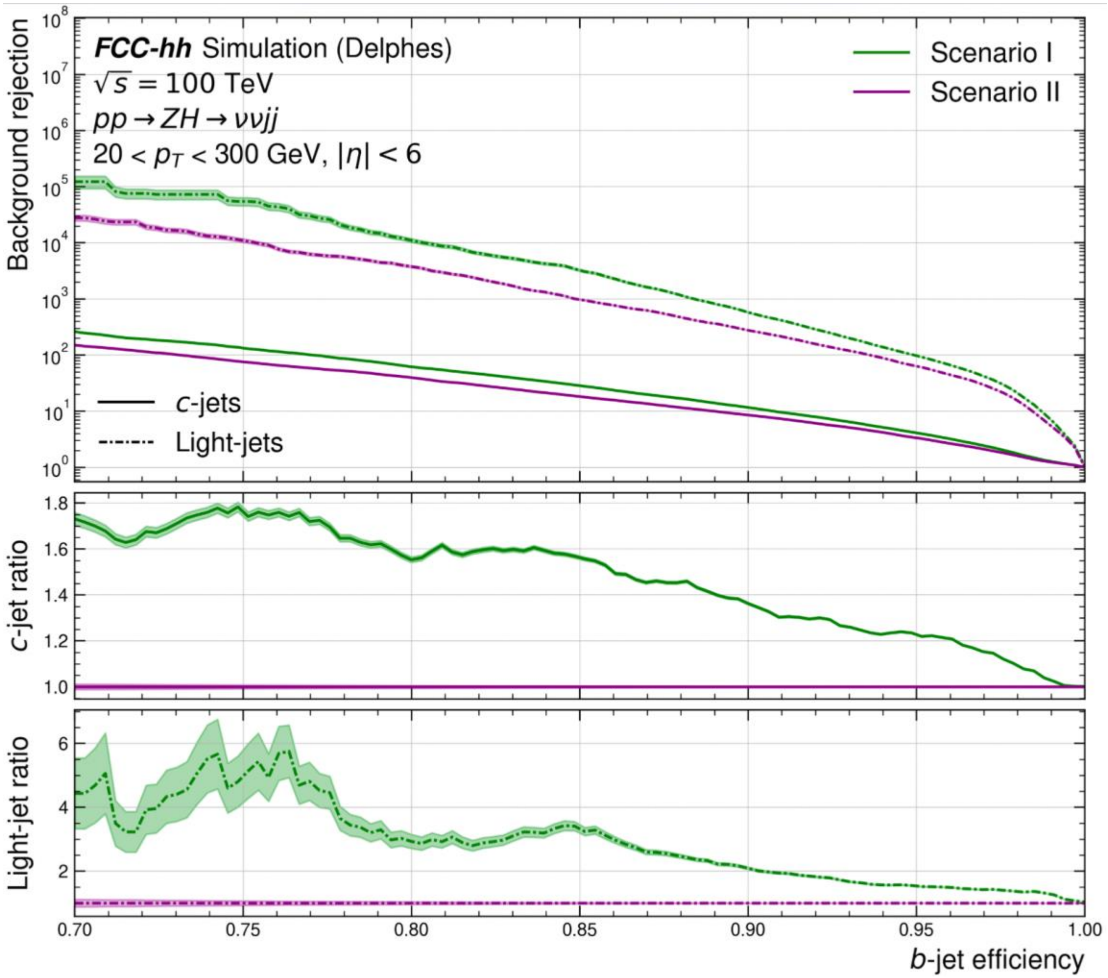
GN2 architecture

Type	Name	GN1	GN2
Hyperparameter	Trainable parameters	0.8M	1.5M
Hyperparameter	Learning rate	$1e{-}3$	OneCycle LRS (max LR $4e{-}5$)
Hyperparameter	GNN Layers	3	6
Hyperparameter	Attention Heads	2	8
Hyperparameter	Embed. dim	128	192
Architectural	Attention type	GATv2	ScaledDotProduct
Architectural	Dense update	No	Yes (dim 256)
Architectural	Separate value projection	No	Yes
Architectural	LayerNorm + Dropout	No	Yes
Inputs	Num. training jets	30M	192M

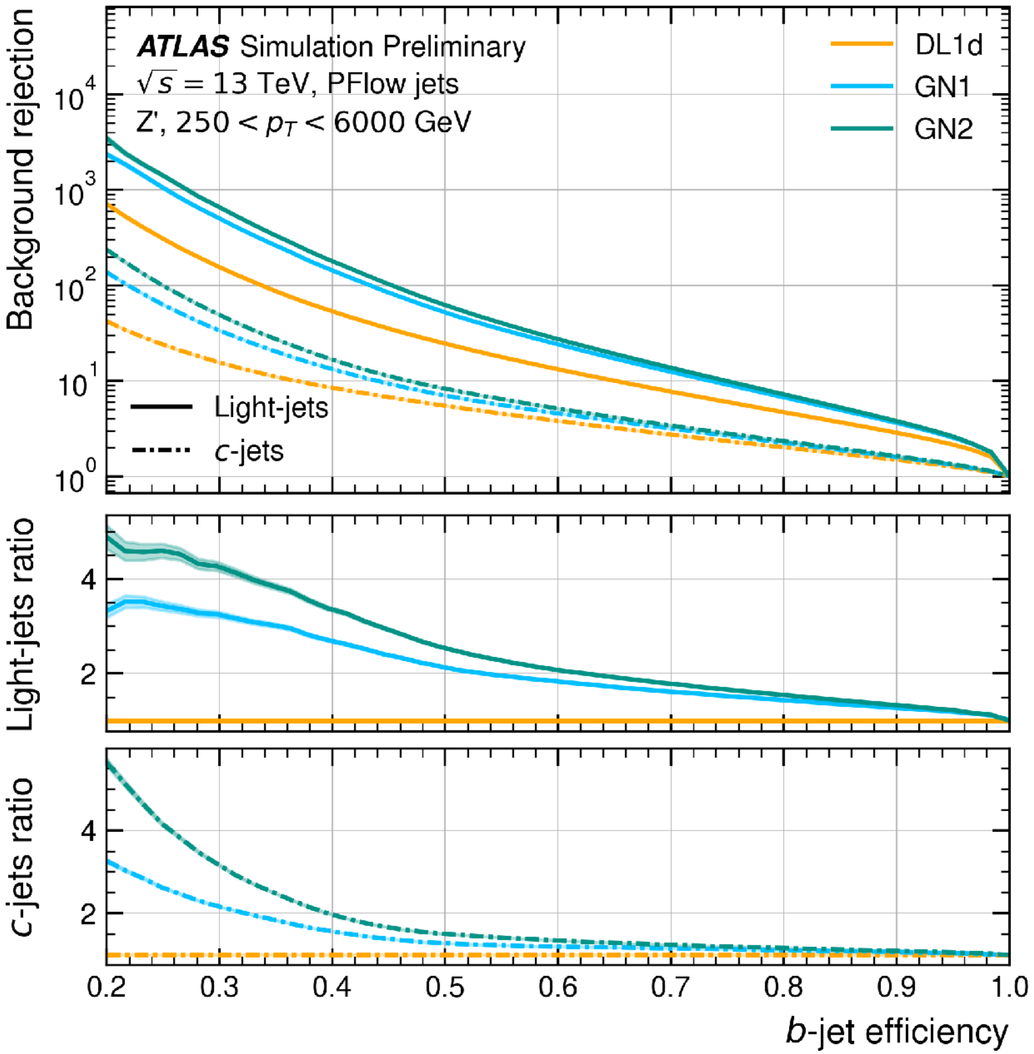
[FTAG-2023-01]

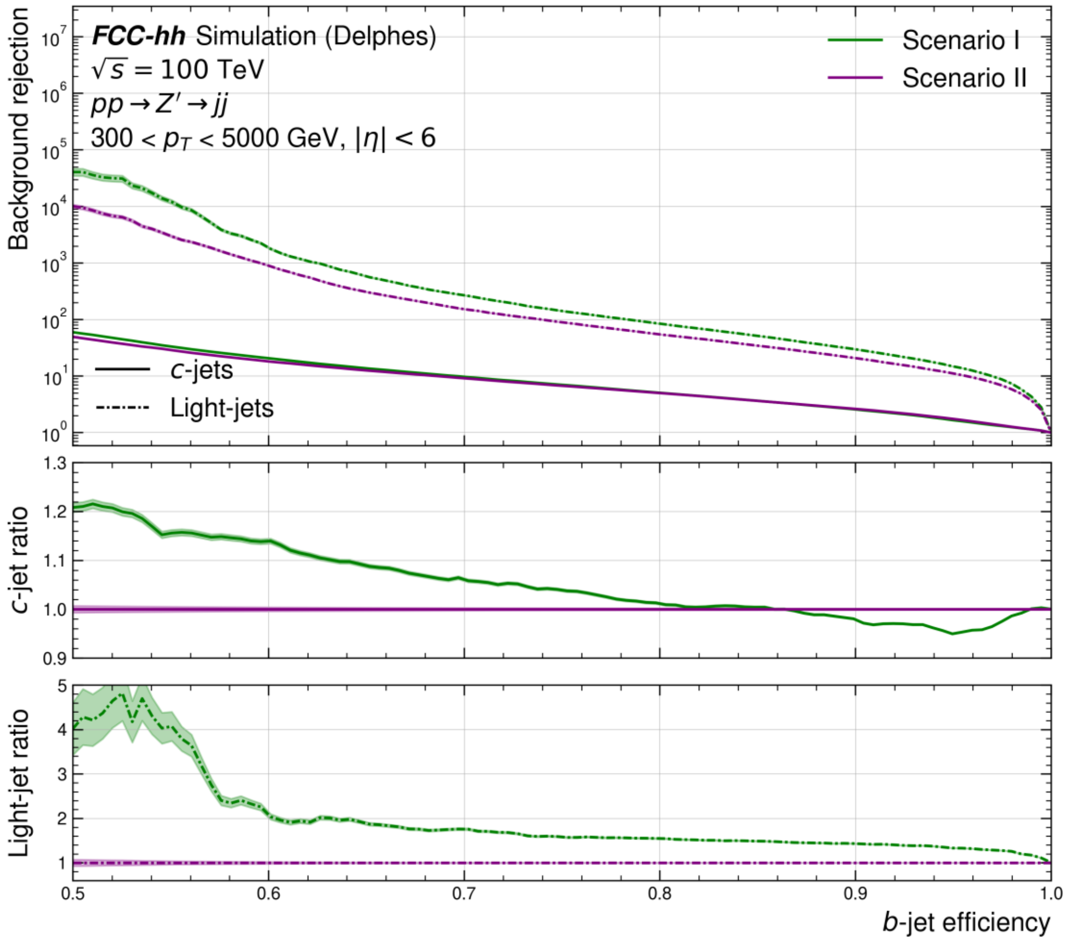
GN2 @ ATLAS vs FCC-hh





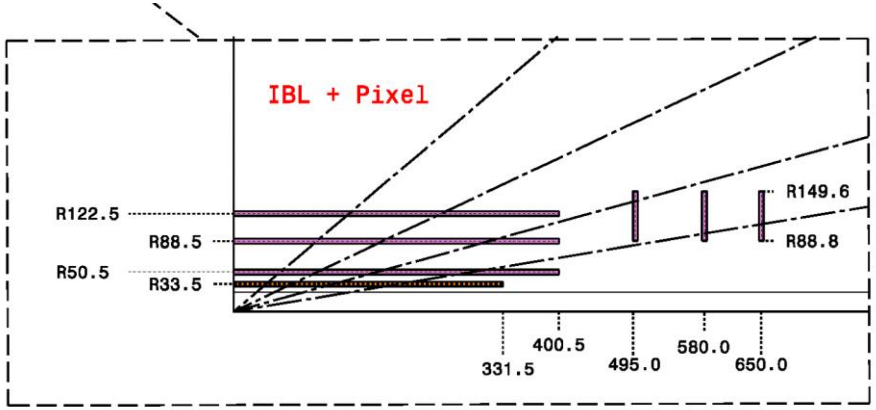
GN2 @ ATLAS vs FCC-hh





GN2 @ ATLAS vs FCC-hh

	LHC (IBL)	FCC-hh
Resolution	10µm x 66µm	7µm x 14µm



Detector envelopes (mm)

IBL	31<R<40
Pixel	42.5<R<242
SCT barrel	255<R<549
SCT end-cap	251<R<610
TRT barrel	554<R<1082
TRT end-cap	617<R<1106

Layer no :	1	2	3	4	5	6	Total
Average radius [mm] :	25.00	60.00	100.00	150.00	270.00	400.00	
Radius-min [mm] :	23.28	58.28	98.28	148.28	261.07	391.07	
Radius-max [mm] :	27.47	63.03	102.52	152.26	280.11	409.73	
Z-min [mm] :	-685.0	-820.0	-820.0	-820.0	-820.0	-820.0	
Z-max [mm] :	685.0	820.0	820.0	820.0	820.0	820.0	
Number of rods :	14	16	26	38	34	50	
Number of modules per rod :	20	40	40	40	17	17	
Number of modules :	280	640	1040	1520	578	850	4908
Disk no :	1	2	3	4	5	Total (+Z & -Z)	
Radius-min [mm] :	25.0	25.0	25.0	25.0	25.0		
Radius-max [mm] :	404.0	404.0	404.0	404.0	404.0		
Average Z pos. [mm] :	950.0	1178.5	1462.0	1813.7	2250.0		
Z-min [mm] :	941.1	1169.6	1453.1	1804.8	2241.1		
Z-max [mm] :	958.9	1187.5	1471.0	1822.6	2258.9		
Number of rings :	4	4	4	4	4		
Number of modules per disk :	108	108	108	108	108	1080	
Ring no :	1	2	3	4			
R-min [mm] :	25.0	101.3	198.9	302.6			
R-max [mm] :	104.7	204.6	303.4	404.0			
Number of modules per ring :	12	20	32	44			

Flat layout:

Pixels (inner)	Macro-pixels (middle)	Striplets/Macro-pixels (outer)
$25 \times 50\,\mu m^2$ (1–4th BRL) $25 \times 50\,\mu m^2$ (1st EC ring) $33.3 \times 100\,\mu m^2$ (2nd EC ring) $33.3 \times 400\,\mu m^2$ (3–4th EC ring)	$33.3 \times 400\,\mu m^2$	$33.3\,\mu m \times 50\,mm$ (BRL) $33.3\,\mu m \times 10\,mm$ (EC)

IBL pitch size:
- 50um x 250um
- Resolution of 10um x 66 um

FCChh detector silicon pixel pitch size:
- 25um x 50um
- Resolution of 7um x 14 um