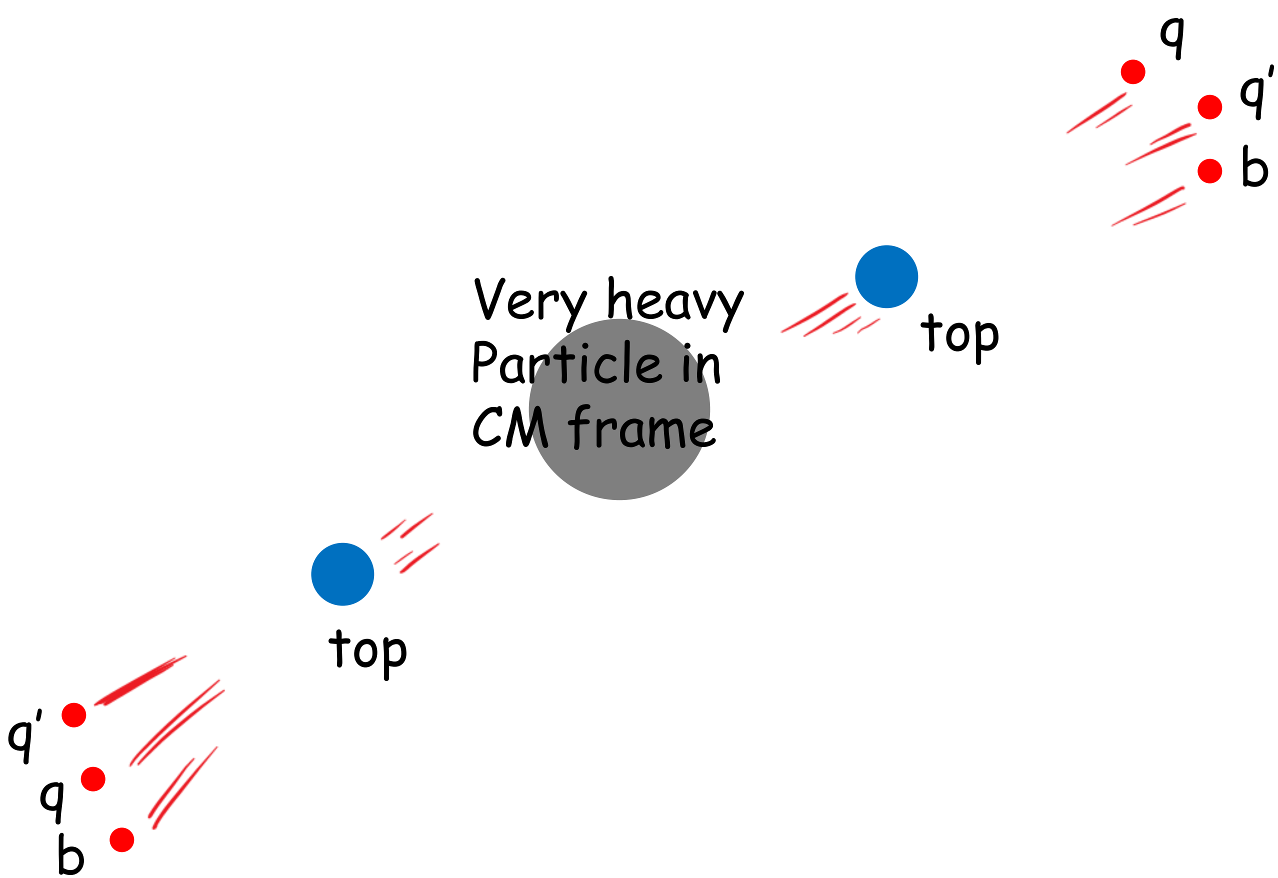
Top-tagging at the Energy Frontier

Minho Son (KAIST)

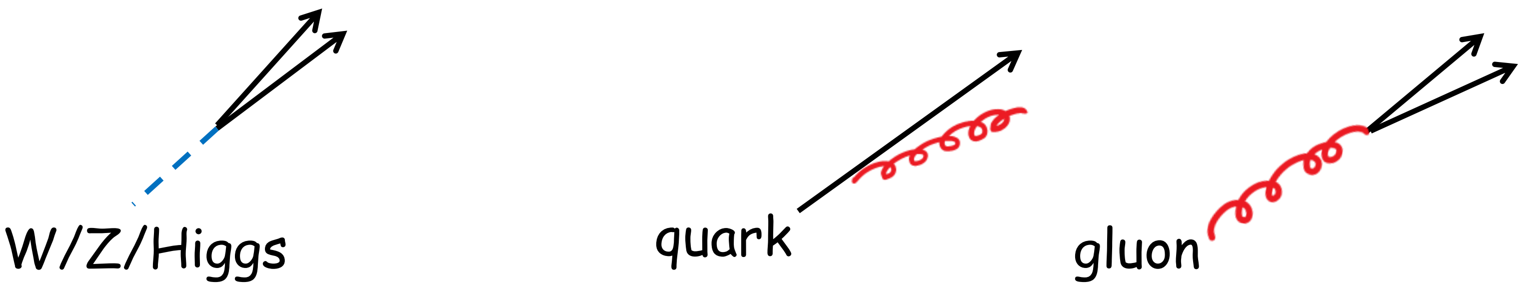
Work in progress with Zhenyu Han and Brock Tweedie

LHC is the first collider in our history which can directly produce TeV-scale particles

- ✓ A large mass gap between electroweak scale and new physics scale produces a lot of boosted W/Z/H/tops

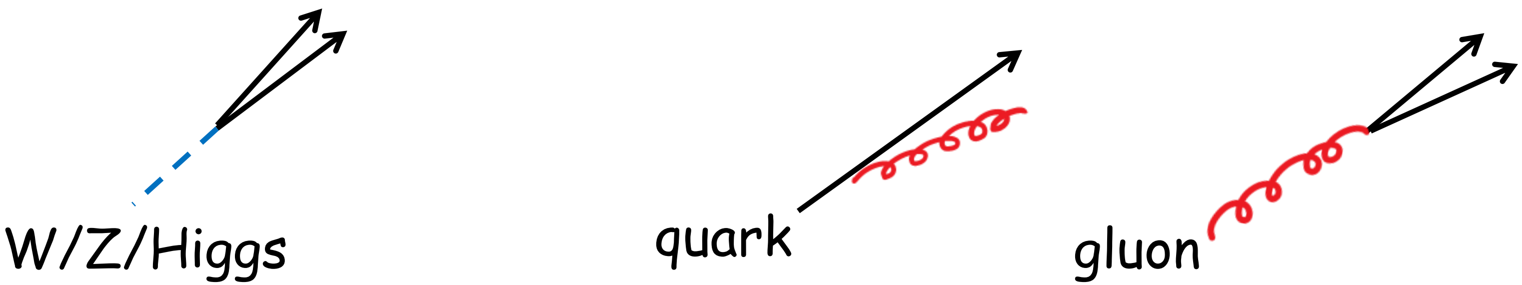


New classification of W/Z/H/top-jets which look similar to QCD-jets



- They look similar, but they are not same in physics
    - ✓ Traditional jet definition is not suitable for disentangling the apparent resemblance

New classification of W/Z/H/top-jets which look similar to QCD-jets



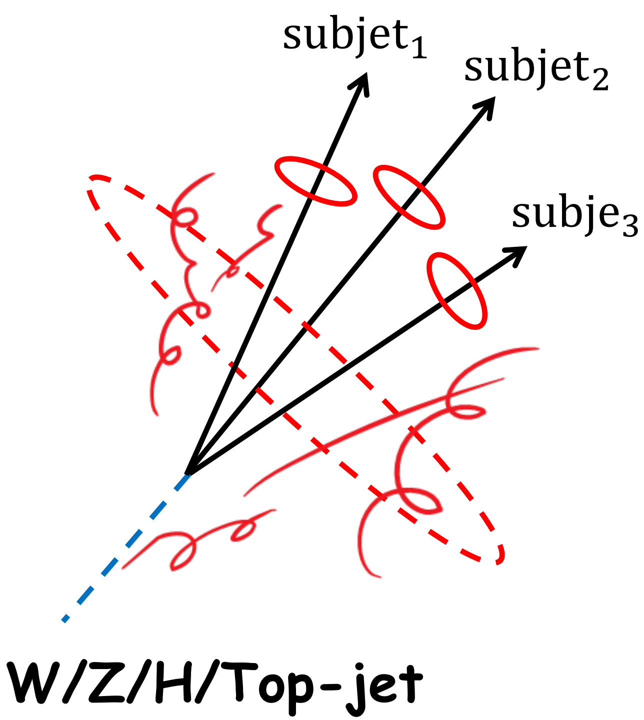
- They look similar, but they are not same in physics
    - ✓ Traditional jet definition is not suitable for disentangling the apparent resemblance ⟶ Emergence of jet-substructure

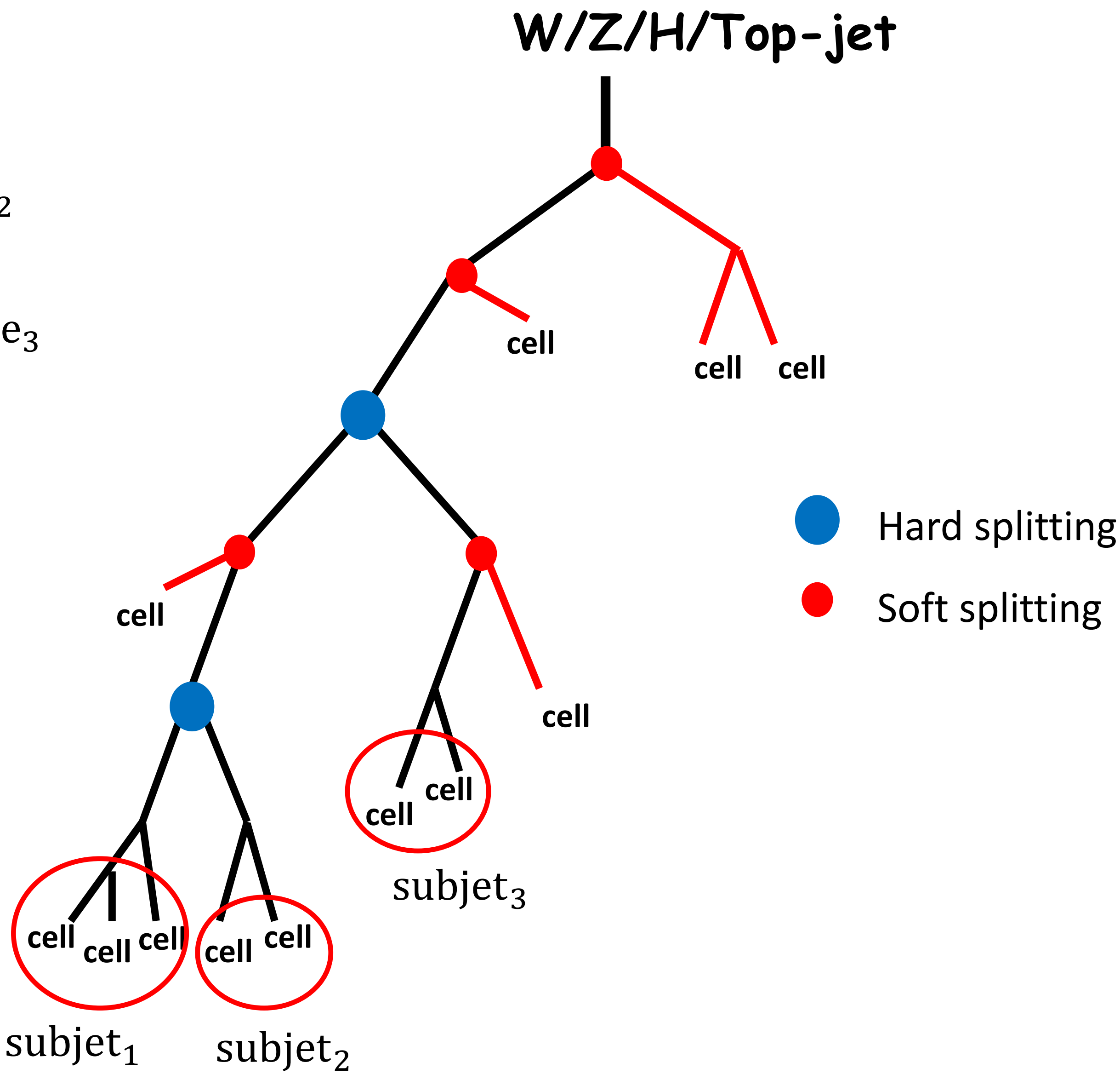
Seymour 1994, Butterworth, Cox, Forshaw 2002
Butterworth, Ellis, Raklev 2007
Butterworth, Davison, Rubin, Salam 2008
….

Jet substructure

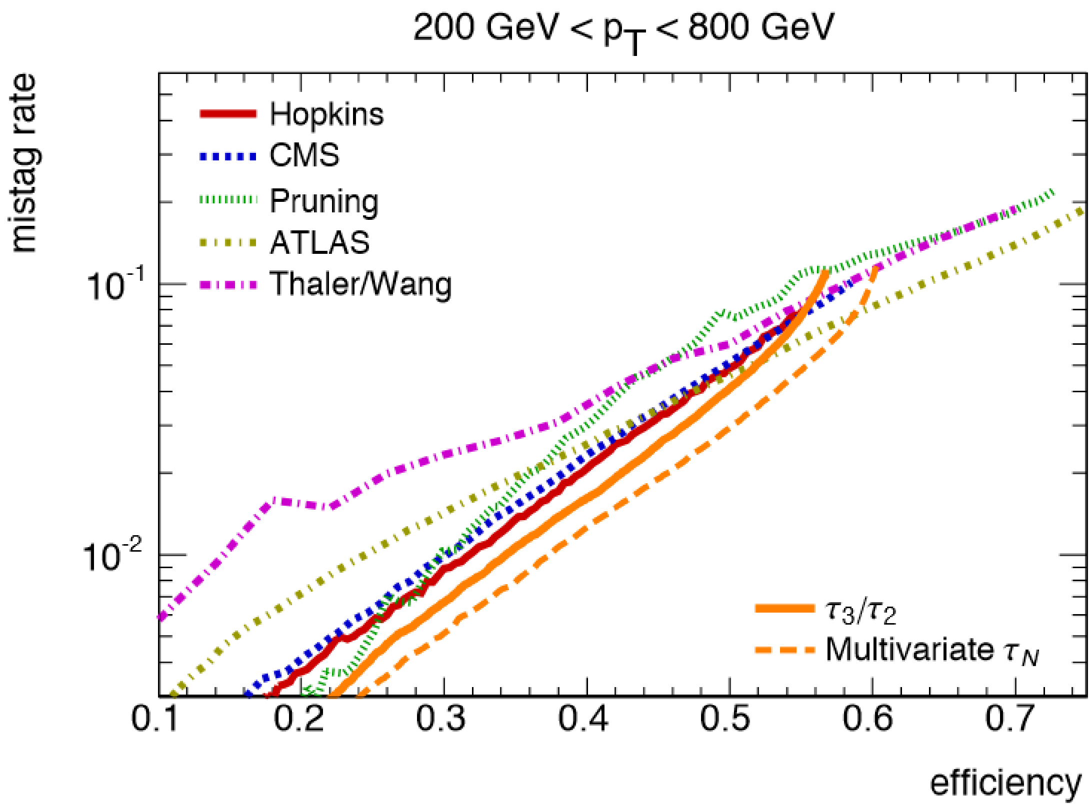
Jet substructure is a right way to organize the hadronic activities in such a way that it correctly reveals true physics

✓ After identifying the region of interest, we‘re looking backward in the jet clustering history to identify hard objects while efficiently filtering out as much QCD contamination as possible





Various Top Taggers in the market



BOOST2010 Proceeding

+ N-subjettiness: Thaler, Tilburg 2011

We are entering into TeV-scale top p_T region, e.g. as the bound on new resonances gets higher, heavier resonances get accessible

How would top tag efficiency evolve with increasing p_T ?

Many challenges arise as tops enter into hyper-boosted regime

✓ We cannot call top "top" unless we can tag it against QCD fakes

Issues on

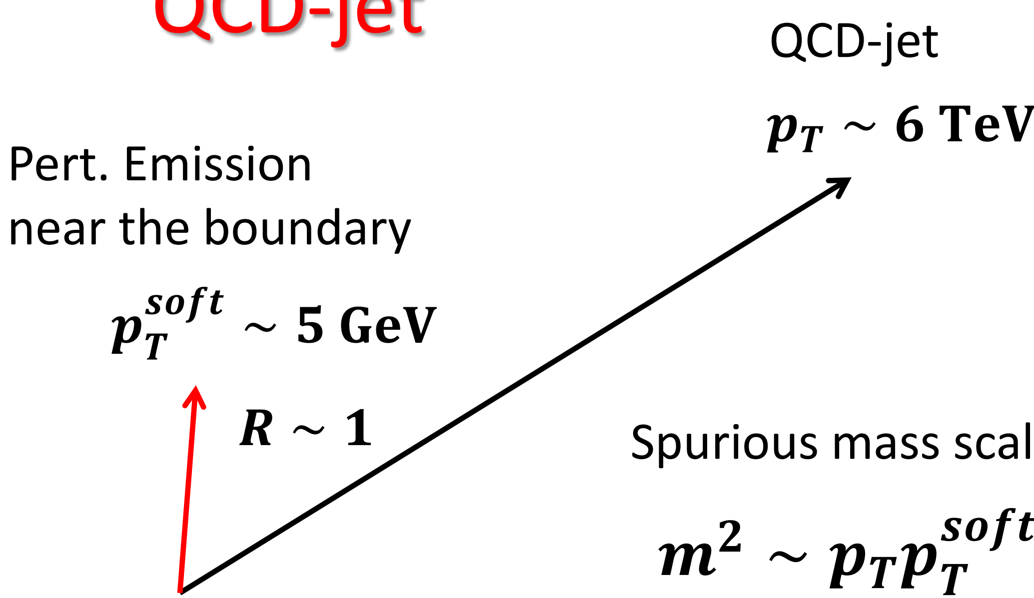
physics

detector

tagger/optimization

I. Instability from soft radiation

QCD-jet



Spurious mass scale

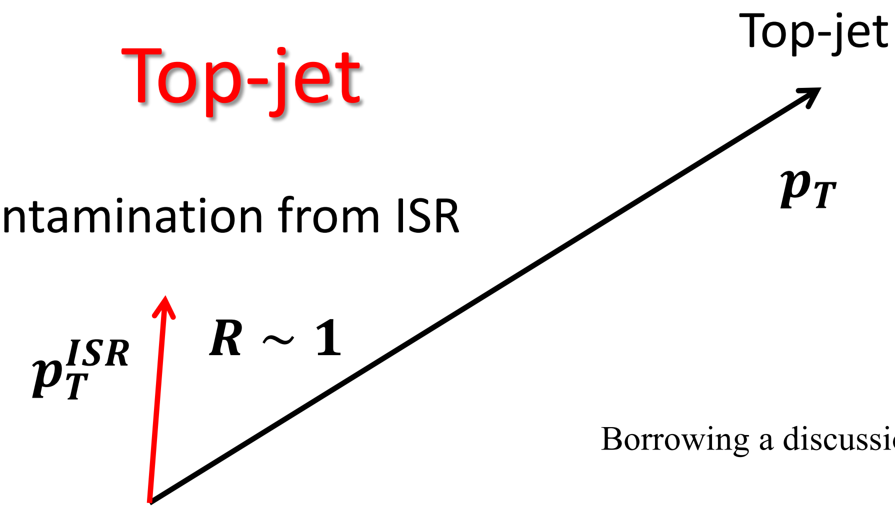
$$m^2 \sim p_T p_T^{soft} R^2 \sim m_{top}^2$$

Borrowing a discussion from Larkoski, Maltoni, Selvaggi 2015

I. Instability from soft radiation

Top-jet

Contamination from ISR



Borrowing a discussion from Larkoski, Maltoni, Selvaggi 2015

Fluctuation of top-jet mass for FIXED cone

$$m^2 \sim m_{top}^2 + p_T p_T^{ISR} R^2 \sim O(1) \times m_{top}^2$$

If $p_T^{ISR} \sim \dfrac{m_{top}^2}{p_T R^2}$

$\sim \mathbf{50\ GeV}$ if $p_T \sim \mathbf{few} \times m_{top}$

Moderately boosted

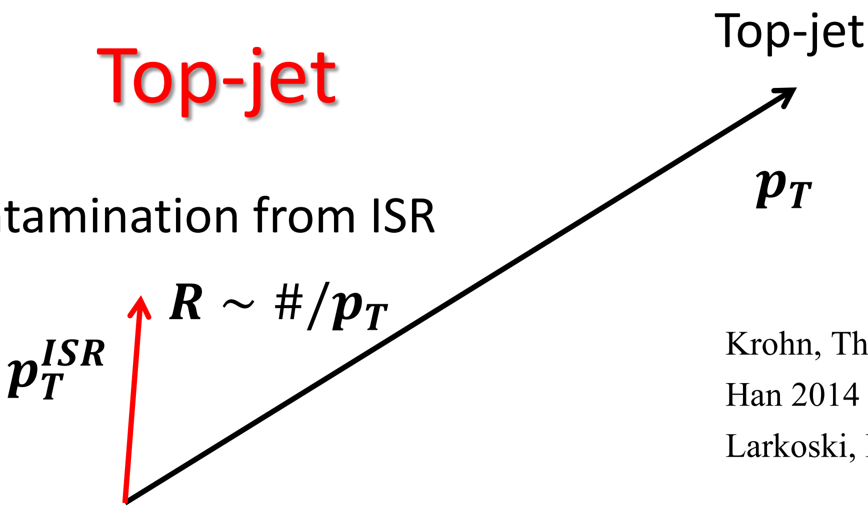
$\sim \mathbf{5\ GeV}$ if $p_T \sim \mathbf{few} \times m_{top} \times \mathbf{10}$

Hyper-boosted

I. Instability from soft radiation vs Shrinking jet size

Top-jet

Contamination from ISR



Krohn, Thaler, Wang 2009
Han 2014
Larkoski, Maltoni, Selvaggi 2015

Fluctuation of top-jet mass for SHRINKING cone

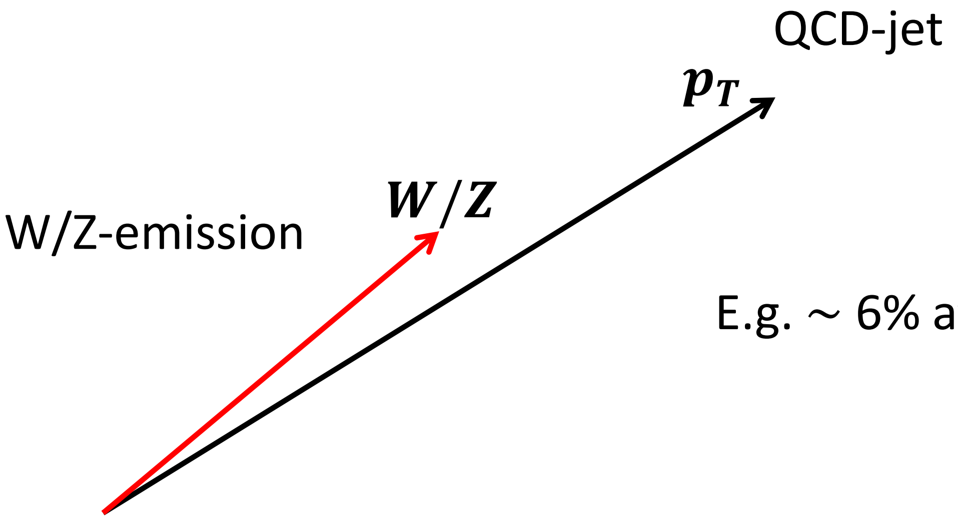
$$m^2 \sim m_{top}^2 + p_T p_T^{ISR}\, R^2 \sim m_{top}^2 \left(1 + \beta_R^2 \frac{p_T^{ISR}}{p_T}\right)$$

With shrinking cone

$$R \sim \beta_R \frac{m_{top}}{p_T}, \qquad \text{e.g. } \beta_R \sim 4$$

II. EW-strahlung at high pT

QCD-jet



E.g. ~ 6% at 5 TeV quark-jet (unpolarized)

Splitting functions (momentum integrated)

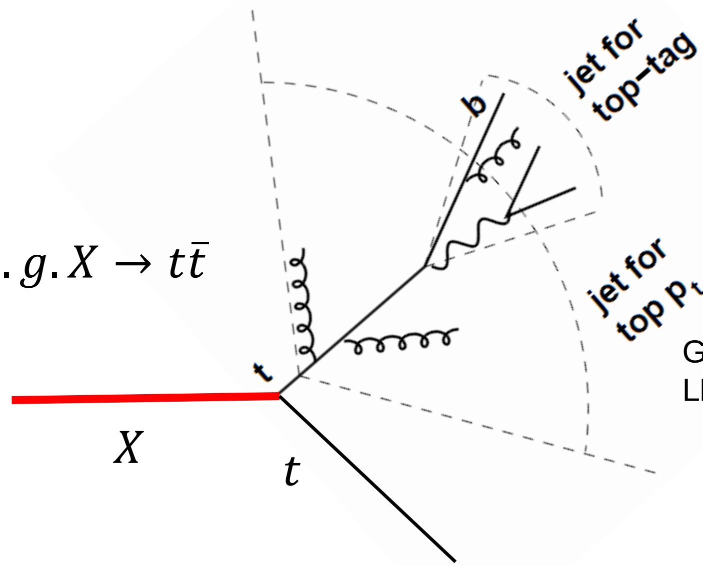
$$P_{q\to W_T} \sim \frac{\alpha_{EW}}{\pi}\frac{1+(1-x)^2}{x}\,\mathrm{Log}\frac{p_T^2}{(1-x)\,m_W^2}$$

$$P_{q\to W_L} \sim \frac{\alpha_{EW}}{\pi}\frac{1-x}{x}$$

- It will be discussed in our paper, but we will not talk about this in this talk

II. Dead cone, FSR, and shrinking cone

$E.g. X \rightarrow t\bar{t}$

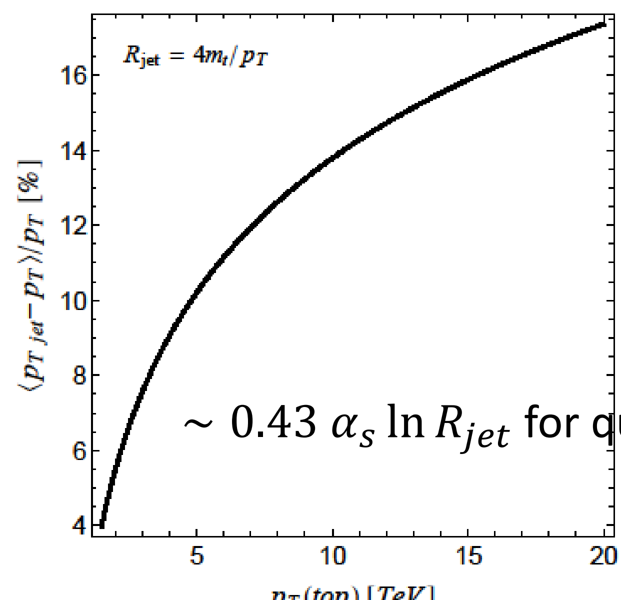


$$R_{dead\ cone} \sim \frac{2\ m_{top}}{p_T}$$

Dead cone captures top decay products , but no radiation from top. Relevant for successful top tagging

G. Salam, talk given at
LHC New Physics Forum, IWH, Heidelberg, 23-26 Feb, 2009

Capturing FSR before decay is important to reconstruct the correct resonance mass where tops were decayed from



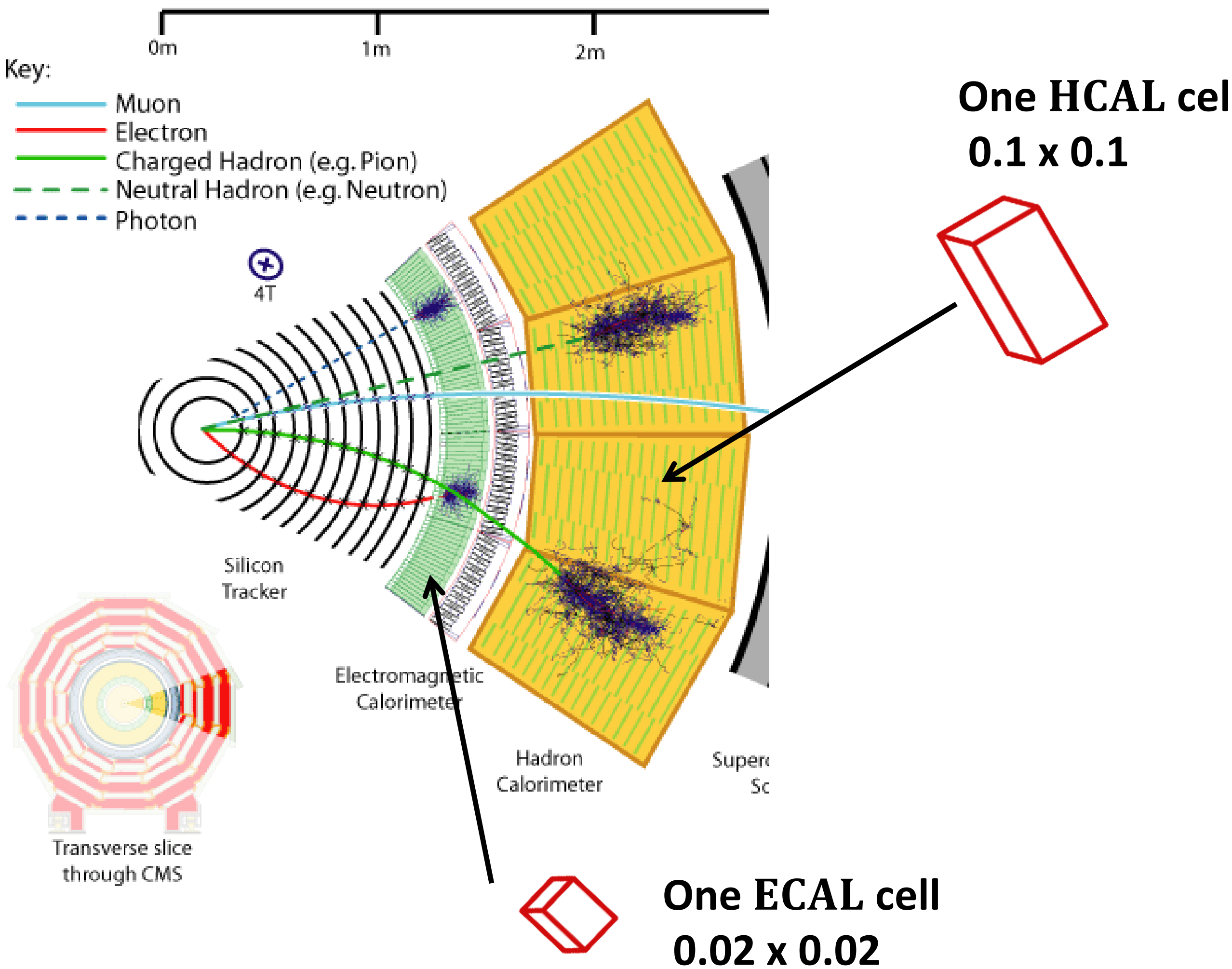
$\sim 0.43\ \alpha_s \ln R_{jet}$ for quark-jet

Salam 2010
'Towards Jetography'

IV. Instrumental challenge

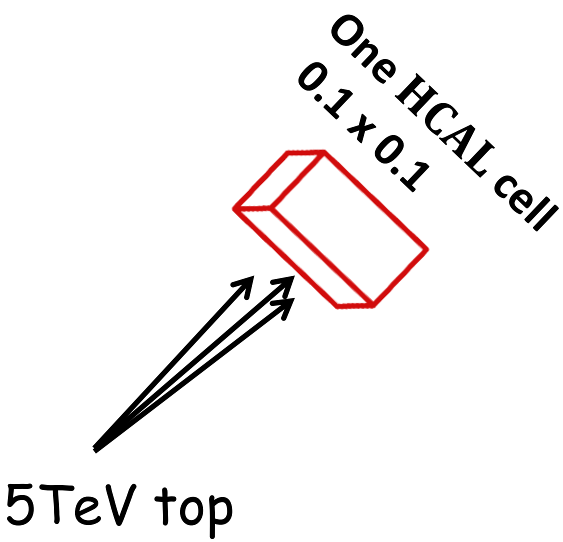
Detector granularity will soon become a big problem.

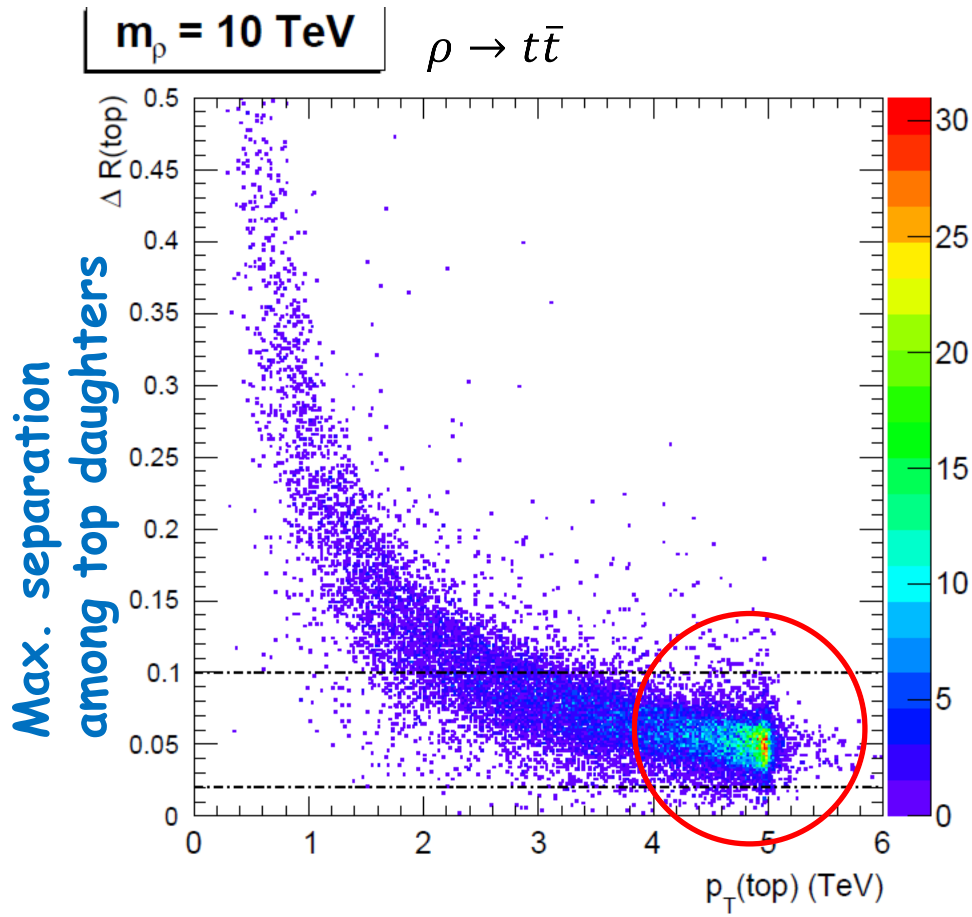
ATLAS/CMS has three layers of main sub-detectors



IV. Instrumental challenge

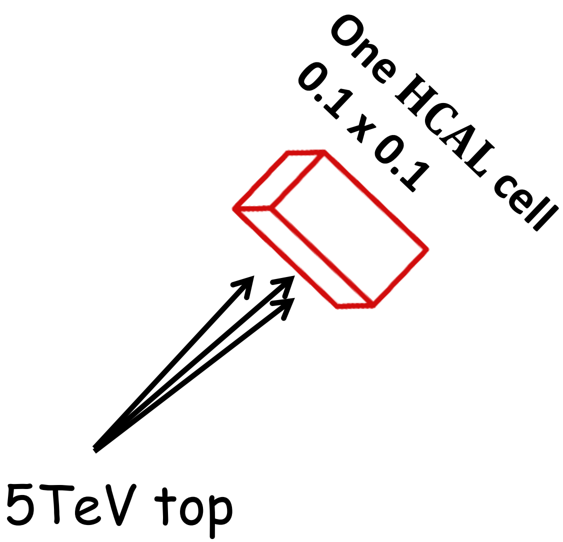
Detector granularity will soon become a big problem.

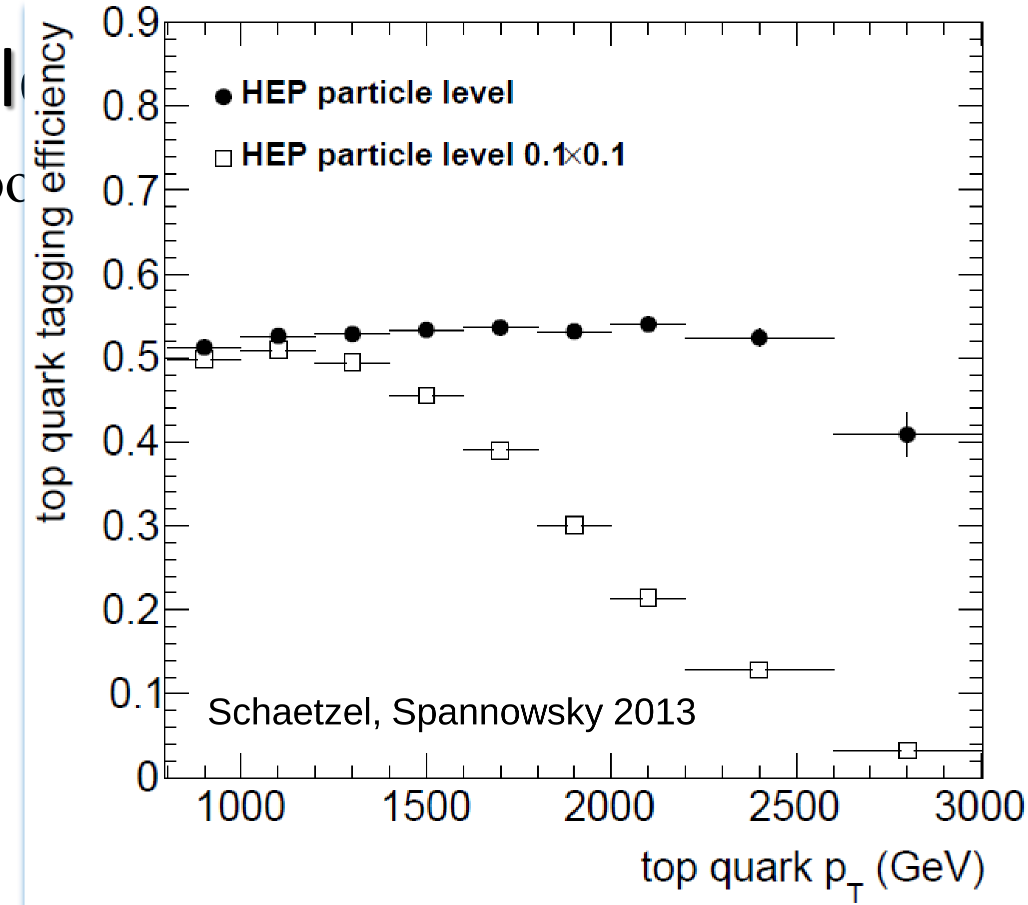


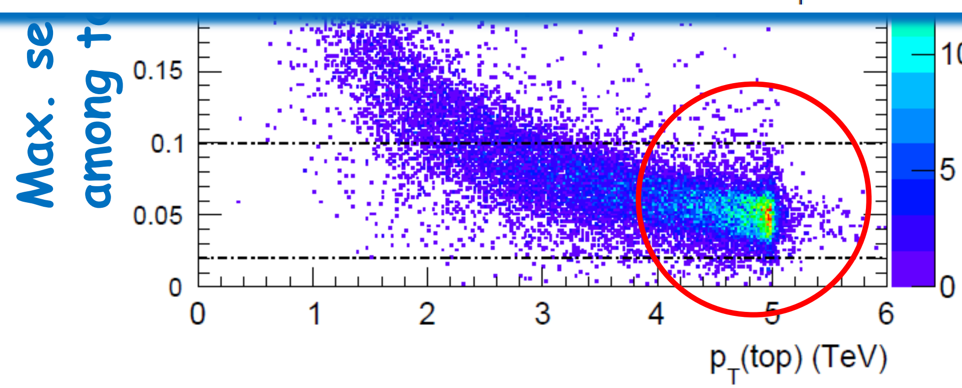


IV. Instrumental chall

Detector granularity will soo

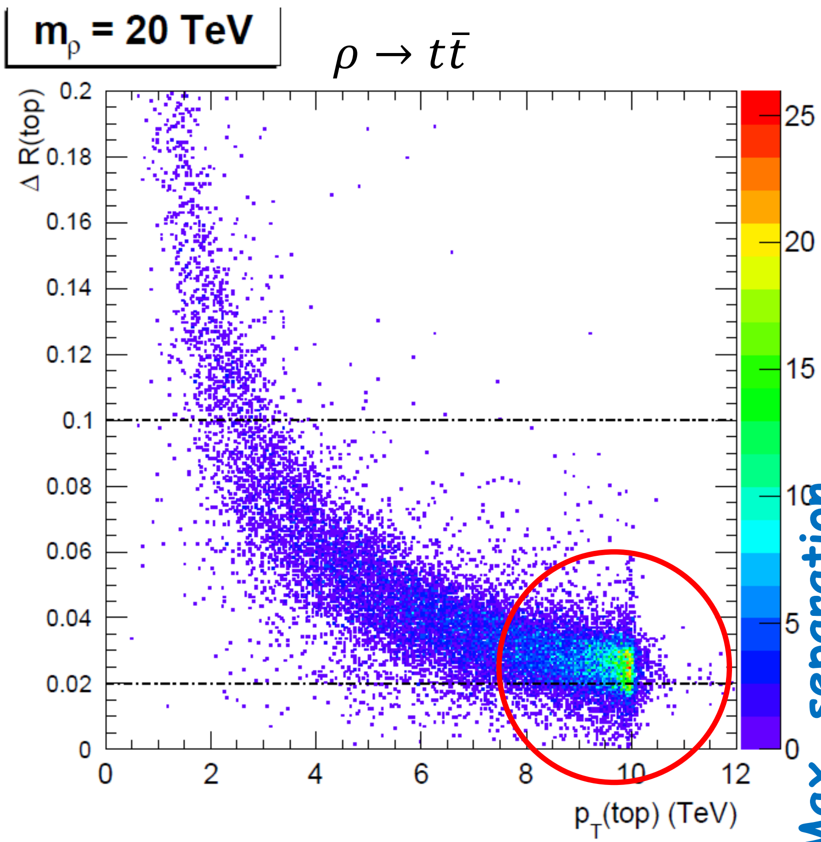




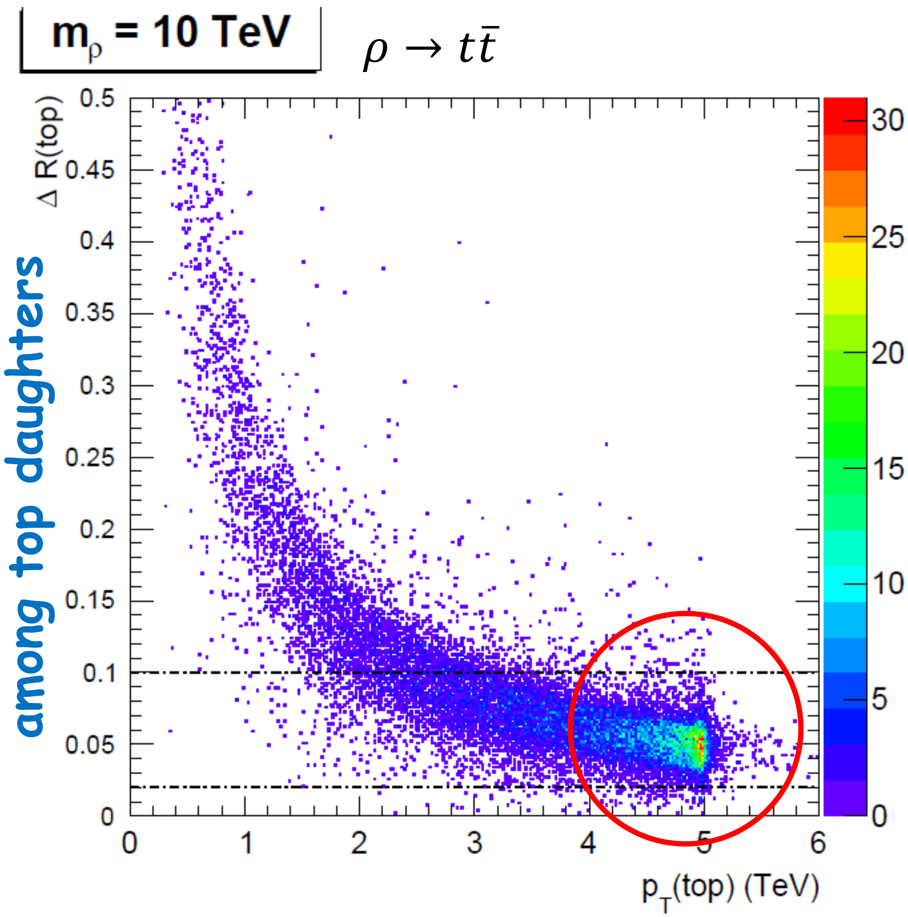


IV. Instrumental challenge

Detector granularity will soon become a big problem.



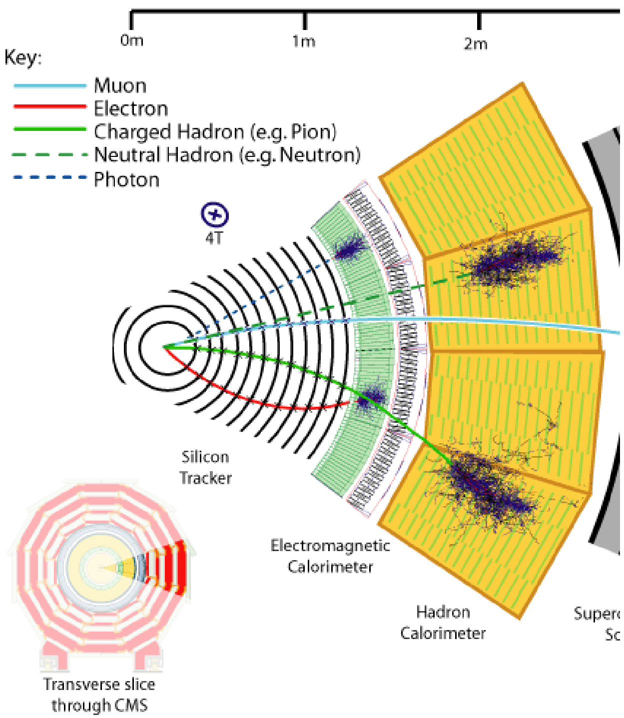
Max. separation among top daughters



IV. Instrumental challenge

Detector granularity will soon become a big problem.

Under 100TeV/14TeV ~ 7x upgrade of CM energy



- Tops from BSM scenarios at the (HL) LHC, FCC

The sensitivity to 3 TeV top at the LHC would expect ~20 TeV tops at 100 TeV collider

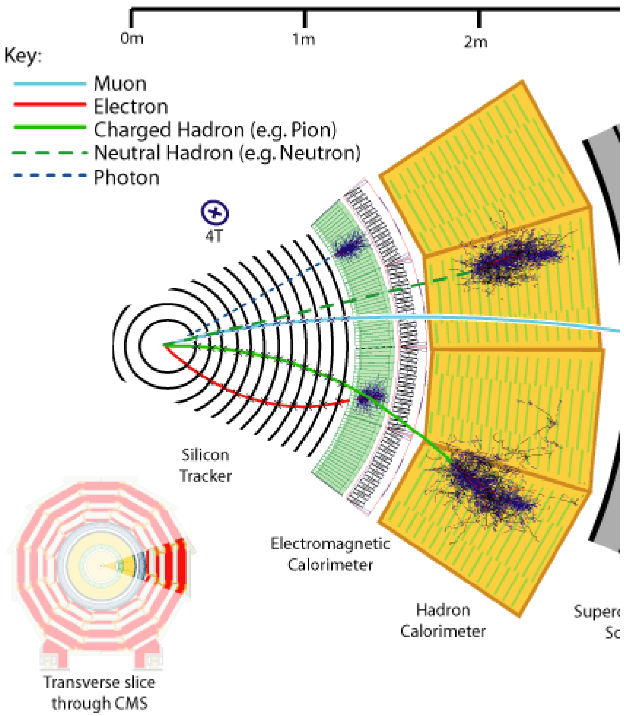
- Current proposal for the future detector

ECAL, HCAL	2x
Tracker	4x

IV. Instrumental challenge

Detector granularity will soon become a big problem.

Under 100TeV/14TeV ~ 7x upgrade of CM energy



- Tops from BSM scenarios at the (HL) LHC, FCC

The sensitivity to 3 TeV top at the LHC would expect ~20 TeV tops at 100 TeV collider

- Current proposal for the future detector

ECAL, HCAL	2x
Tracker	4x

Not scale-invariant!

✓ Unless we come up with a new idea even at the LHC-type detector, we will be out of our business at FCC

Capability at the LHC $+$ **FCC proposal** $\rightarrow$ **Capability at the FCC**

Capability at the LHC $+$ **FCC proposal** $\rightarrow$ **Capability at the FCC**

If 5TeV top can be tagged

Assuming Nx better resolution

* Detector cost is proportional to Volume, i.e. $\sim \mathrm{N}^3$, assuming same materials though

5N TeV top can be tagged

Capability at the LHC + **FCC proposal** → **Capability at the FCC**

If 5TeV top can be tagged

Assuming Nx better resolution

* Detector cost is proportional to Volume, i.e. $\sim \mathrm{N}^3$, assuming same materials though

5N TeV top can be tagged

Understanding our current detector better is a KEY-ingredient to predict our future capability

Existing Literature

Katz, MS, Tweedie, Spethmann 2011, 2012
Snowmass 2013
Schaetzel, Spannowsky 2013
CMS PAS JME-14-002 2014
Spannowsky, Stoll 2015
Larkoski, Maltoni, Selvaggi 2015
….

* Listed only studies on W/Z/H/tops-taggers

Outline

We will focus on JHU TopTagger + N-subjettiness

1. Optimize JHU TopTagger + N-subjettiness at particle level

   We will newly show that N-subjettiness is not just an alternative to other top-taggers, but it adds a new information to improve top/gluon discrimination

2. Introduce various detector models

   We will illustrate how one can combine information, scattered in here and there in sub-detectors, to extract a meaningful result

3. Optimize JHU TopTagger + N-subjettiness in various detector models

   This step will establish the "robustness of shape variables vs declustering variables against different detector configurations"

JHU TopTagger with CMS type cuts

Top-jet size

$$R_{jet} = \beta_R \times \frac{m_t}{p_T}$$

At each branch,

$$\delta_p = \frac{p_T(j_i)}{p_T(top\ jet)}$$

$$\delta_r = \beta_r \times \frac{m_t}{p_T}$$: min angular dist.

j_{top}

↓ First iteration

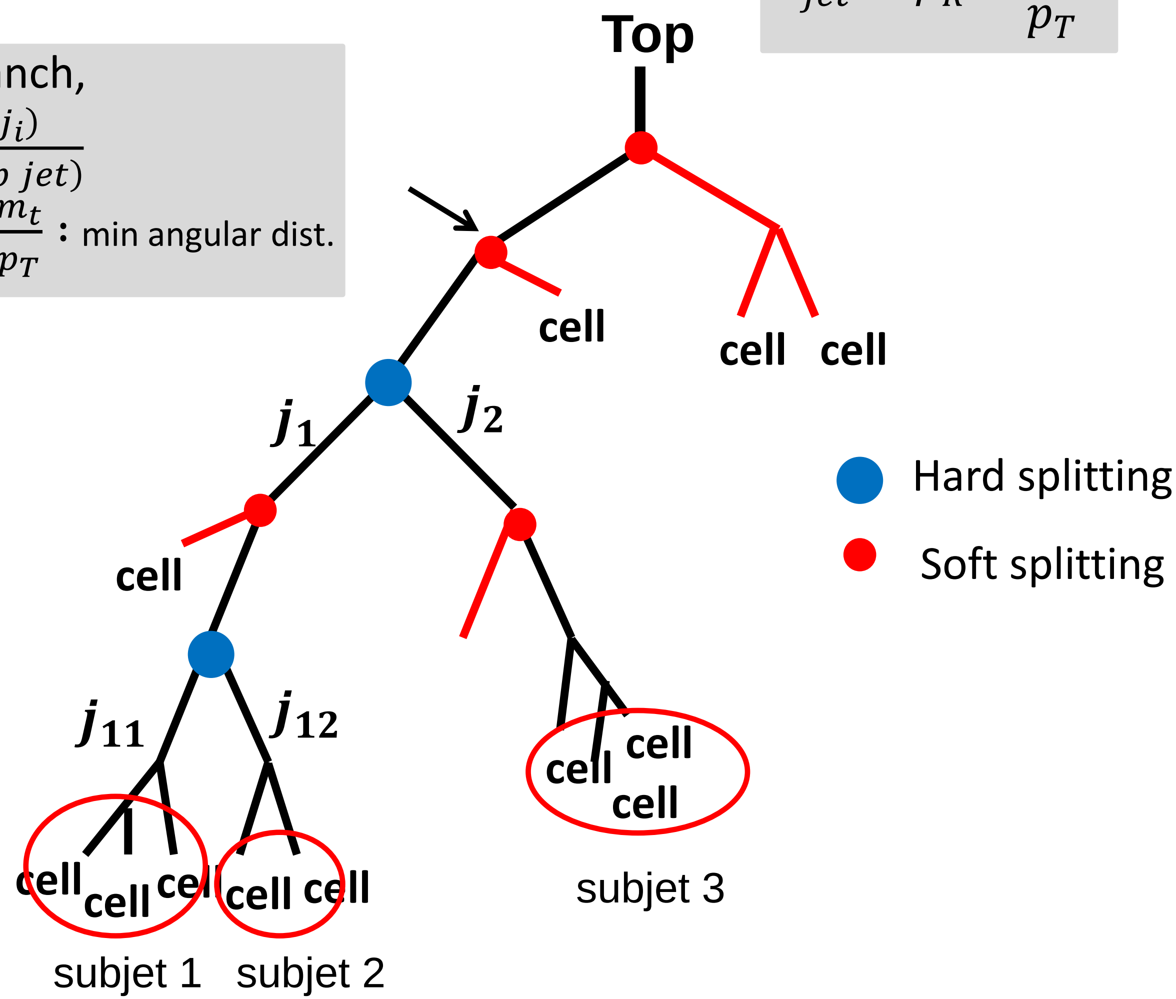
j_1 j_2

↓ Second iteration

j_{11} j_{12} j_{21} j_{22}

⇩

3 or 4 subjets



Cuts on two variables: m_{min}, m_{top}

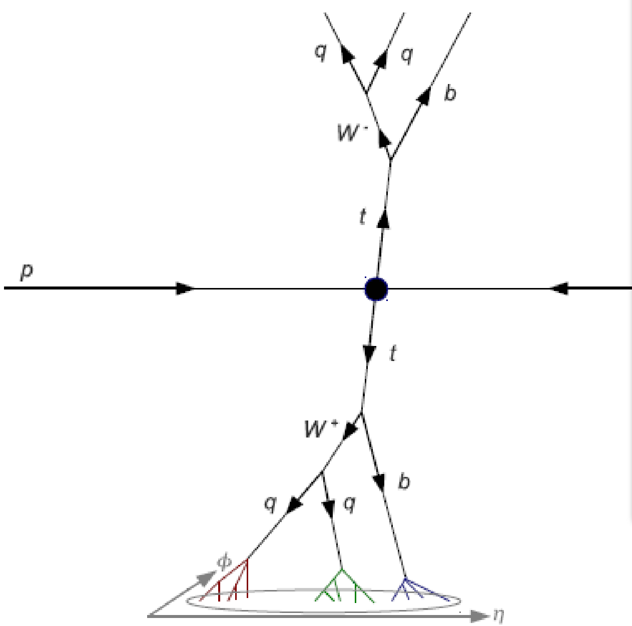
* Instead of m_W and $\cos\theta$ in the original JHUTopTagger

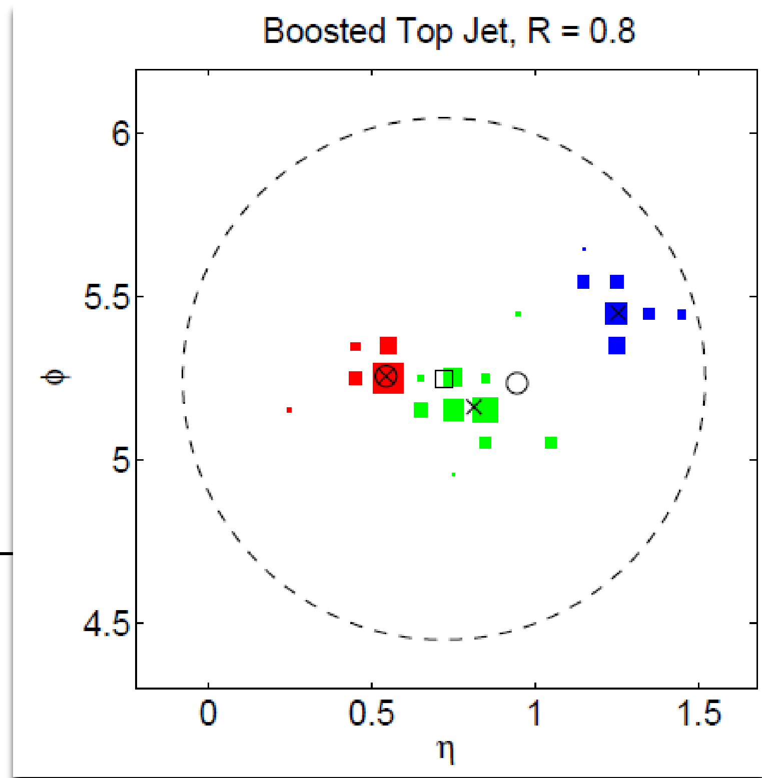
N-subjettiness

- exploits radiation pattern around N subjet axes

$$\tau_N \equiv \frac{\min\left[\sum p_T(i)\,\min\{\Delta R(i,\widehat{j_1}),\ldots,\Delta R(i,\widehat{j_N})\}\right]}{\sum p_T(i)\;R}$$

Thaler, Tilburg JHEP 1103



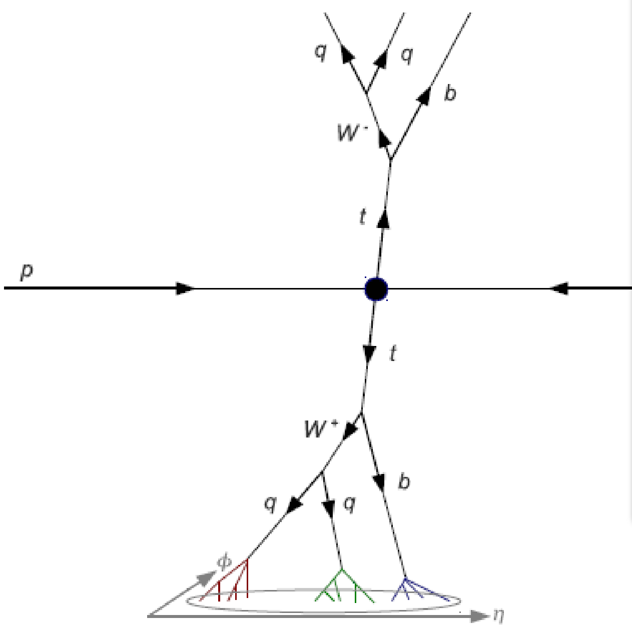


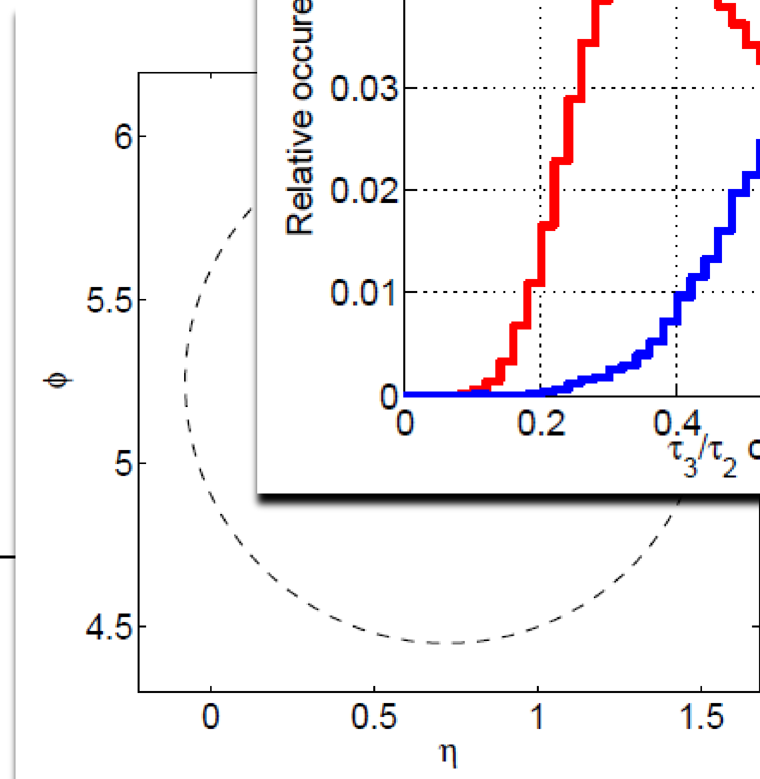
Top: three localized energy clusters

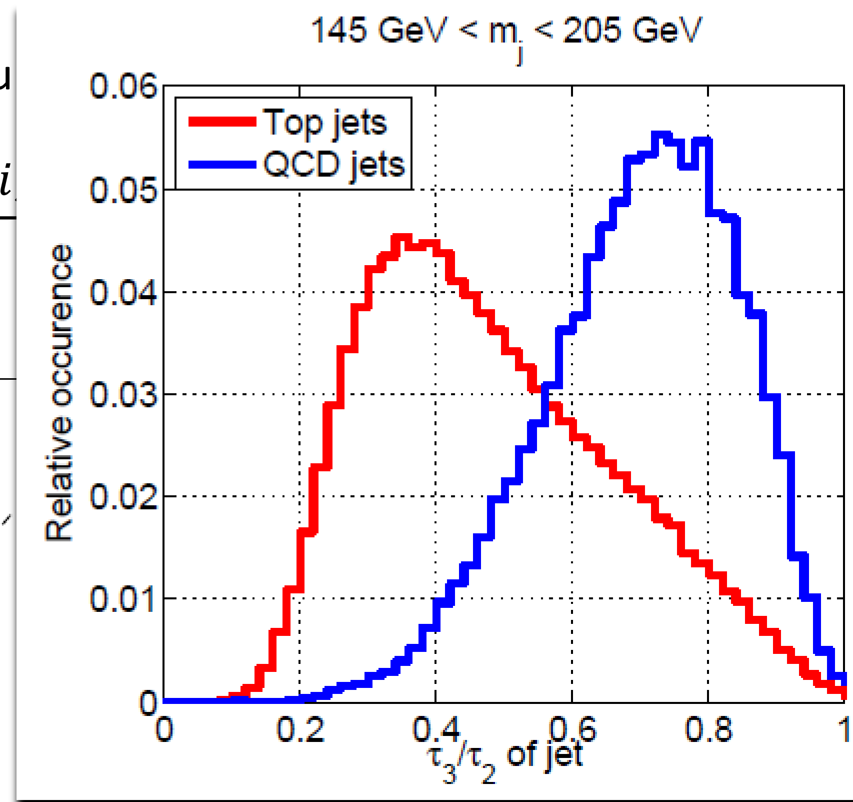
N-subjettiness

- exploits radiation pattern around N su

$$\tau_N \equiv \frac{\min\left[\sum p_T(i) \min\{\Delta R(i,\hat{j}_1), \ldots, \Delta R(i\right.}{\sum p_T(i)\, R}$$







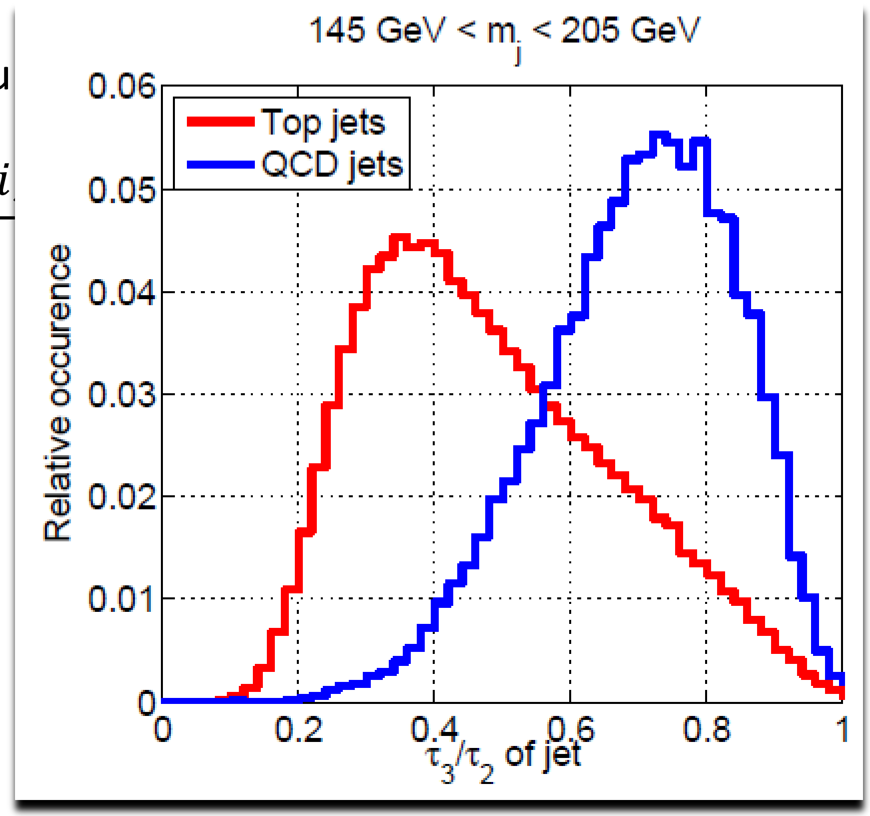
Thaler, Tilburg JHEP 1103

Top: three localized energy clusters

N-subjettiness

- exploits radiation pattern around N su

$$\tau_N \equiv \frac{\min\left[\sum p_T(i)\,\min\{\Delta R(i,\hat{j}_1),\ \ldots.,\ \Delta R(i\right.}{\sum p_T(i)\ R}$$



Thaler, Tilburg JHEP 1103

✓ N-subjettiness is qualitatively different from other top taggers based on mass/pT-drops and it has been introduced as an alternative for top tagger

✓ We newly observe that combining other top taggers with N-subjettiness can actually give $O(1)$ improvement in top/gluon discrimination

Optimization

JHU Top-tagger with CMS-type cuts & N-subjettiness

Clustering/declustering/cut parameter

$R_{\mathrm{Anti-kt}} = 1.0$

$R_{jet} \equiv \beta_R \times \frac{m_t}{p_T}$: Shrinking jet-cone size

δ_p: pT asymmetry cut , $\delta_r \equiv \beta_r \times \frac{m_t}{p_T}$: min angular separation

m_{min}: min jet pair mass m_{top}: reco- top mass $\tau_{32} \equiv \tau_3/\tau_2$: N-subjettiness

Optimization over seven parameters

$$\text{Tag/mistag Rate} \equiv \frac{\text{\# survived to the end}}{\text{\# generated with 1\% pT window}}$$

Signal: continuum $t\bar{t} \to \mu + jets$ Quark/gluon: $qZ \to q(\nu\bar{\nu})$, $gZ \to g(\nu\bar{\nu})$

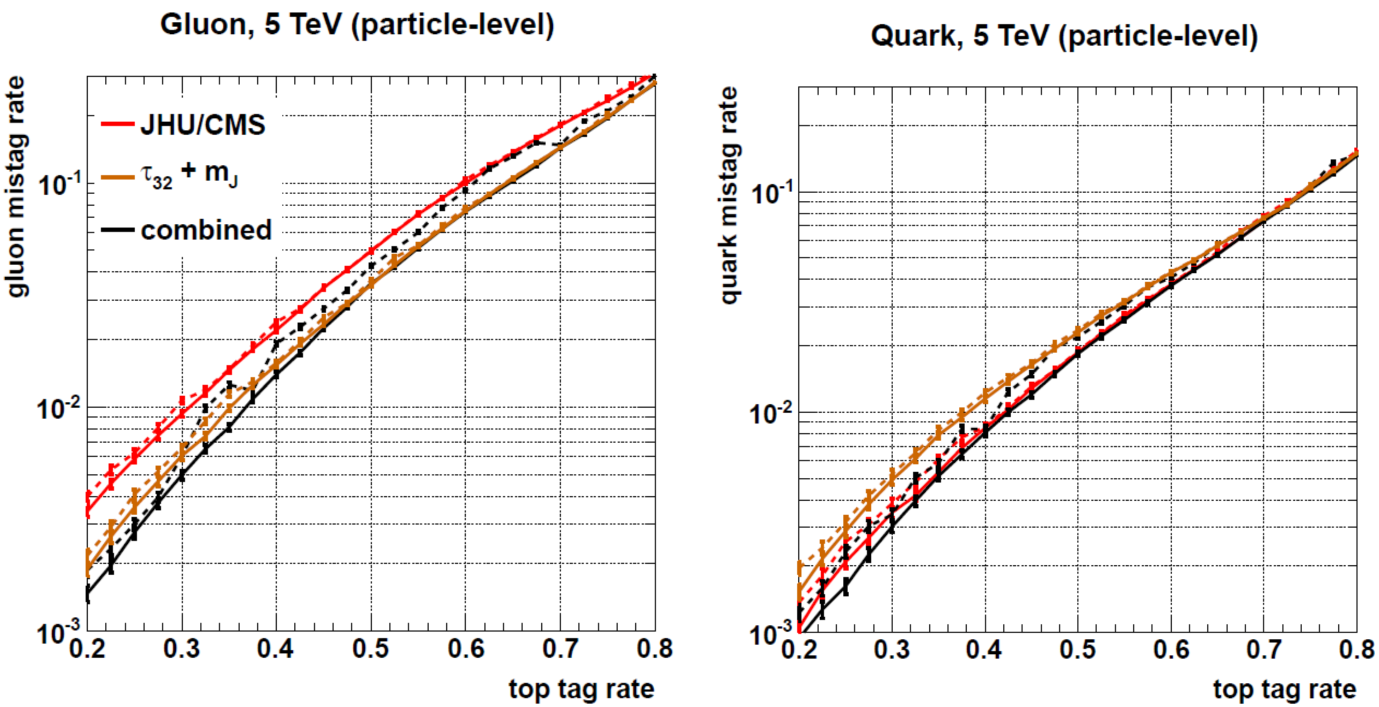
: samples are restricted to $|\eta| < 1.0$, $p_T = [p_T - 1\%, p_T + 1\%]$ GeV

Top/gluon/quark discrimination at particle level

Flavor dependent optimizations:

With gluon-optimized parameters

With quark-optimized parameters



$\beta_R \sim 4,\ \beta_r \sim 0.7,\ \delta_r \sim 0.03$ for relevant tag efficiencies

- Simultaneous optimizations of the quark and gluon jets by JHU/CMS are possible
- N-subjettiness adds extra discriminating power for gluon-jets, not quark-jets

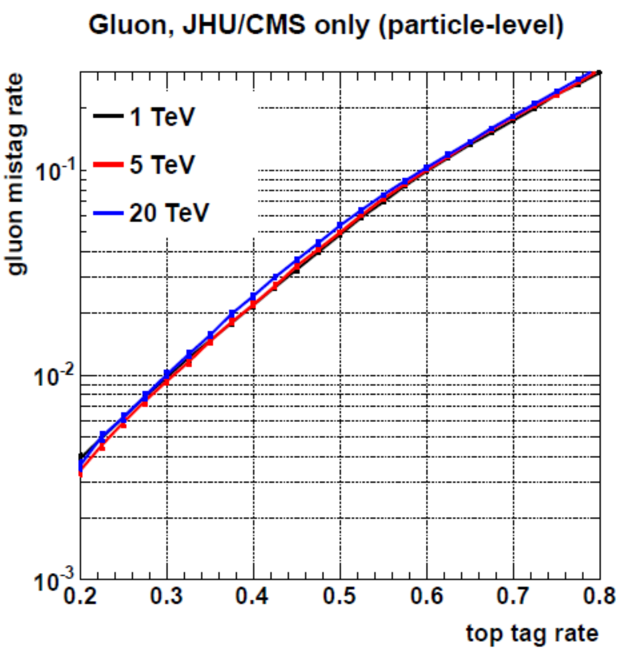
- Gluon-mistag with quark-optimized parameters gets worsen (not conclusive though)

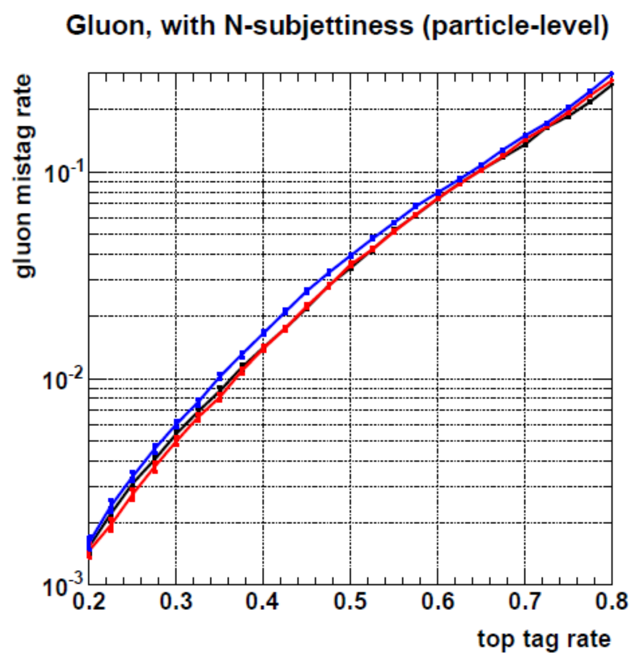
Top/gluon discrimination at particle level

p_T-dependent optimizations on top/gluon-jets:

Two separate optimizations:

JHU with CMS-type cuts without vs with N-subjettiness





- Nearly scale-invariant!
- N-subjettiness adds extra discriminating power for gluon-jets
- Optimized parameters are roughly unchanged, e.g. optimized β_R and β_r stay fixed, simple $\sim 1/p_T$ scaling works

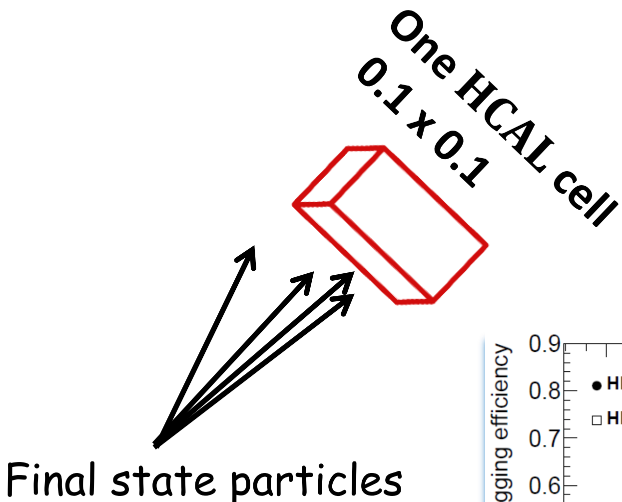
Introducing **detector effect**

What is a good detector model?

It is the one that minimally breaks the ‘scale invariance’ and brings the result back to our expectation at the ‘particle-level’

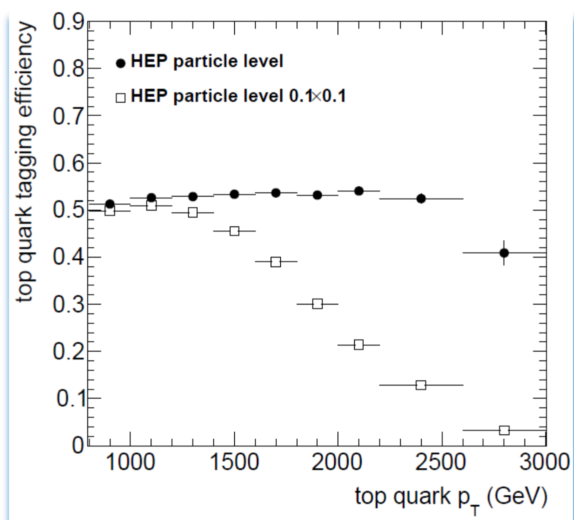
Introducing detector effect

While the real detectors are insanely complicated, our toy detector model would catch the leading effects. However, we are aiming to be as close to the reality as possible



HCAL cell size serves as a cut-off in many pheno- studies

❑ Raw HCAL



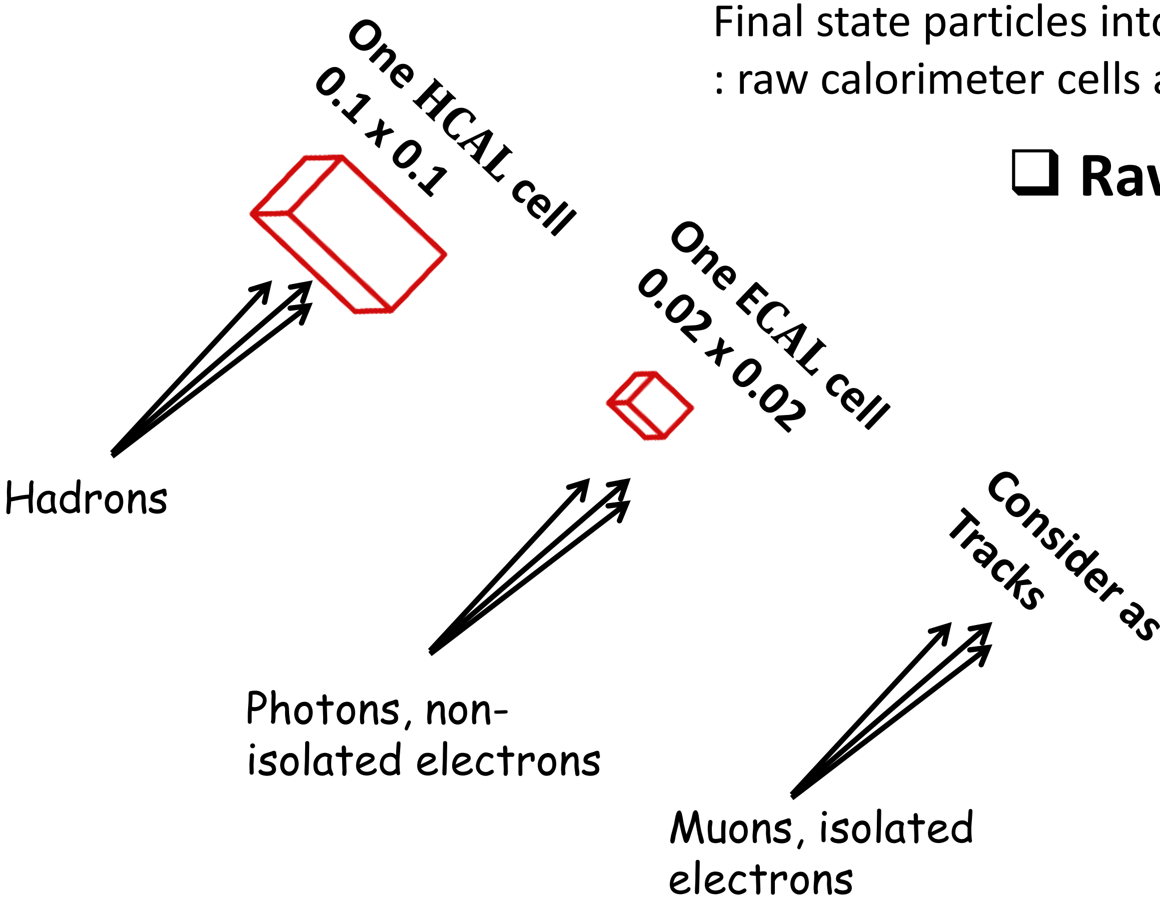
Schaetzel, Spannowsky 2013

Introducing detector effect

While the real detectors are insanely complicated, our toy detector model would catch the leading effects. However, we are aiming to be as close to the reality as possible

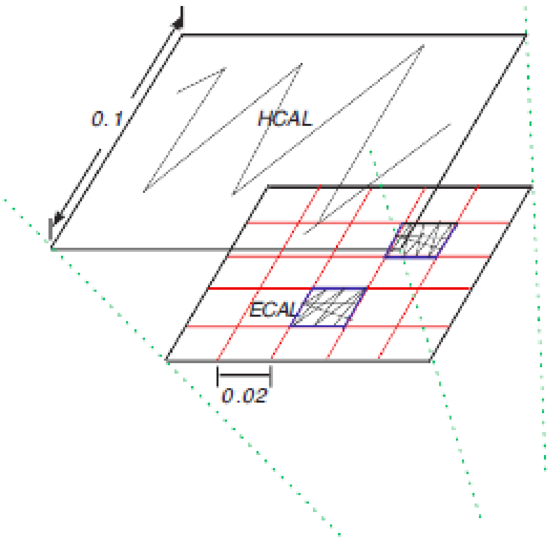
Final state particles into three different sub-detectors : raw calorimeter cells as inputs for jet clustering

- ❑ **Raw ECAL & HCAL**



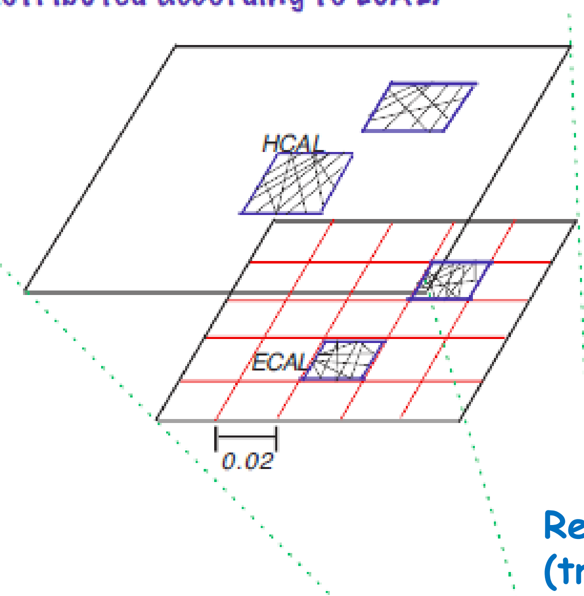
Toy detector models

Cartoon picture of our toy detector model



ECAL is rescaled to match
full ECAL + HCAL energy
(i.e. HCAL is re-distributed according to ECAL)

ECAL to trace
jet energy flow



Repeat with trackers
(track flow, particle flow)

✓ Use every bit of information from three layers of sub-detectors

Katz, MS, Spethmann,Tweedie, 2011, 2012 (See Appendices of 1010.5253/1204.0525)

Toy detector models

Combining information is not unique

Toy detector models

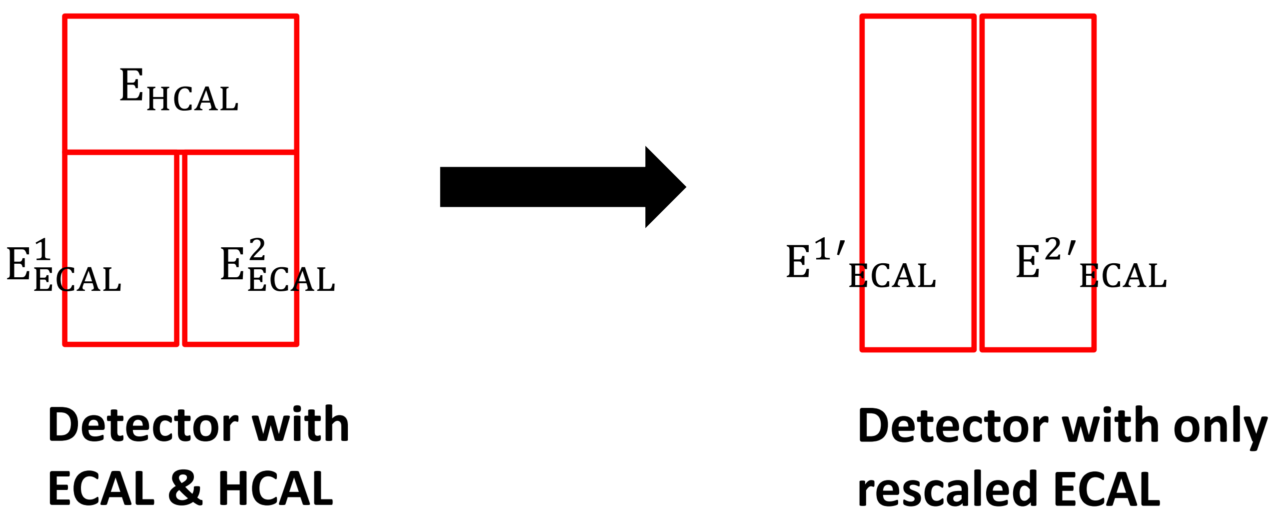
Combining information is not unique

- **EM-flow**

ECALs are locally rescaled to the energy of the full calorimeter, and HCAL cells discarded

Katz, MS, Spethmann, Tweedie 2011, 2012
(See Appendices of 1010.5253/1204.0525)

Rescale ECAL cells by $\frac{E_{ECAL}+E_{HCAL}}{E_{ECAL}}$



* Rescaled ECAL cells are input for the jet clustering

Toy detector models

Combining information is not unique

- **EM-flow** ECALs are locally rescaled to the energy of the full calorimeter, and HCAL cells discarded

  Katz, MS, Spethmann, Tweedie 2011, 2012

  Rescale ECAL cells by $\frac{E_{ECAL}+E_{HCAL}}{E_{ECAL}}$

- **Track-flow** Similarly rescale tracks by $\frac{E_{ECAL}+E_{HCAL}}{E_{tracks}}$

  Schatzel, Spannowsky 2014

  Larkoski, Maltoni, Selvaggi 2015

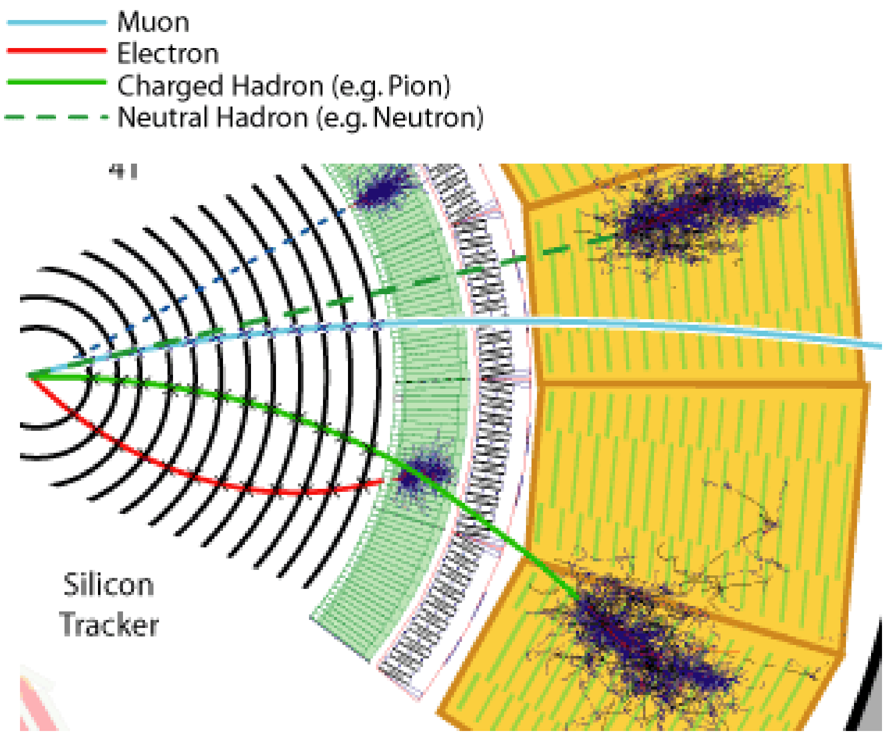
- **Particle-flow** Rescale tracks by $\frac{E_{HCAL}}{E_{tracks}}$ and leave E_{ECAL} as-is

* PERFECT tracking efficiency is assumed. Reality is worse than this perfect case

Two crucial detector effects added to be more realistic

1. Energy-smearing into nearby calorimeter cells

2. Hadrons deposit their energies in ECAL cells



Unlike the situation in this cartoon, hadrons have O(1) chance to leave their energies (e.g. via Nuclear interaction) in ECAL before reaching HCAL.

O(20%) of jet energy becomes absorbed in the ECAL in this manner

Two crucial detector effects added to be more realistic

1. Energy-smearing into nearby calorimeter cells

2. Hadrons deposit their energies in ECAL cells

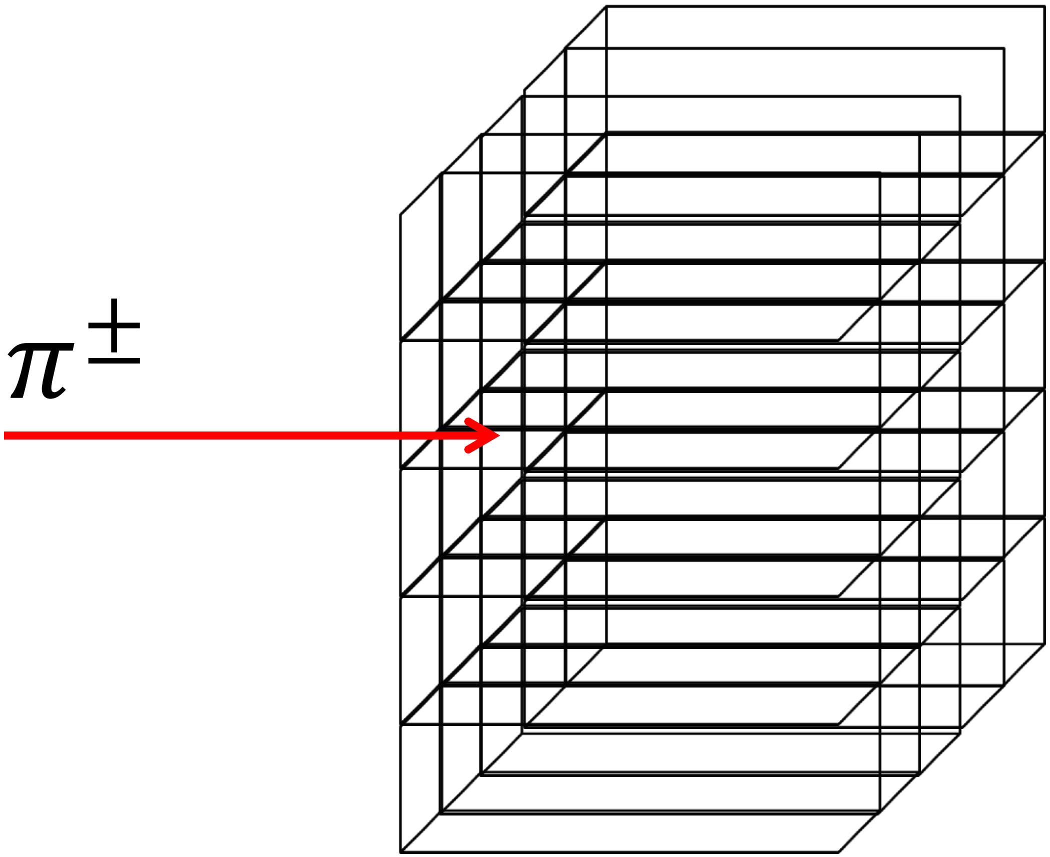
GEANT

- ECAL smearing pattern/hadron-energy-deposit-in-ECAL will be simulated with GEANT4 whereas HCAL smearing pattern will be done by simple ansatz

Upgraded version of MS, Spethmann, Tweedie 2012 (See Appendix of 1204.0525)

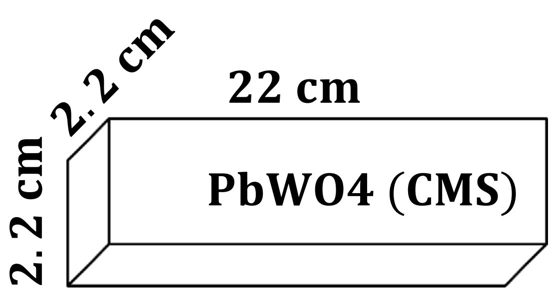
Energy smearing into nearby ECAL cells

- ✓ The most important ingredient in our detector model



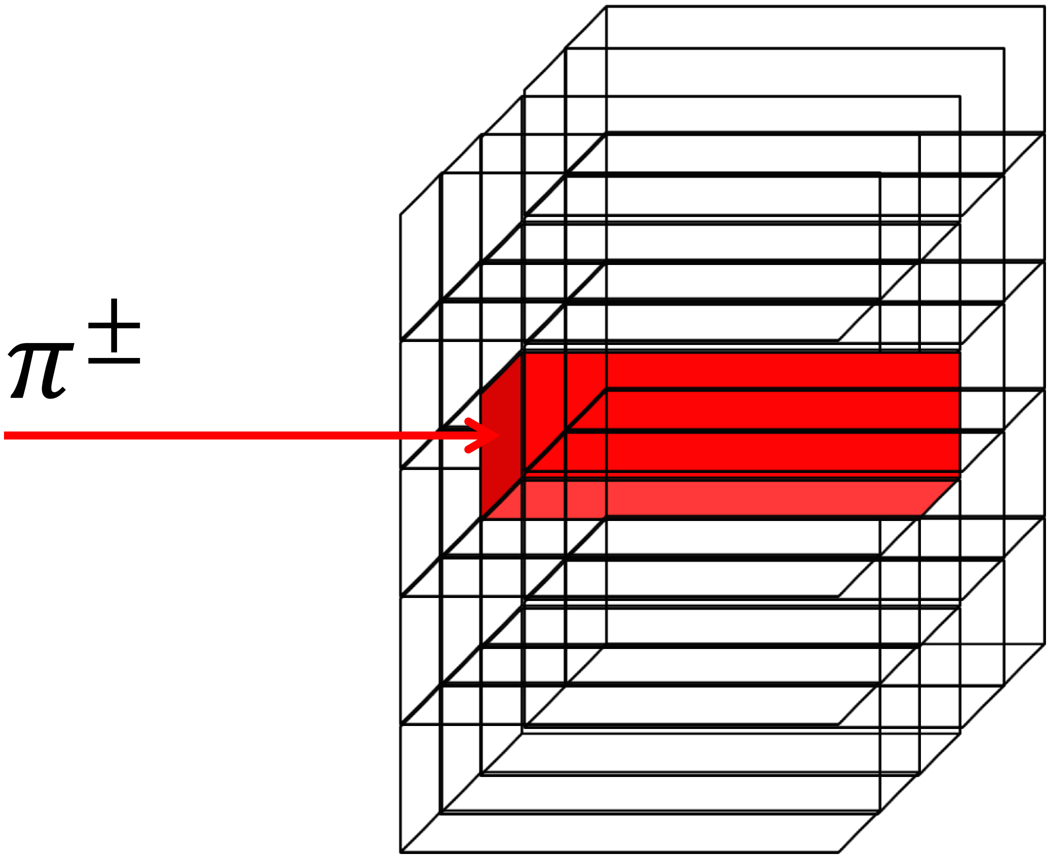
A particle hitting ECAL cell

5 x 5 ECAL cells underneath one HCAL cell



Energy smearing into nearby ECAL cells

Ideal situation

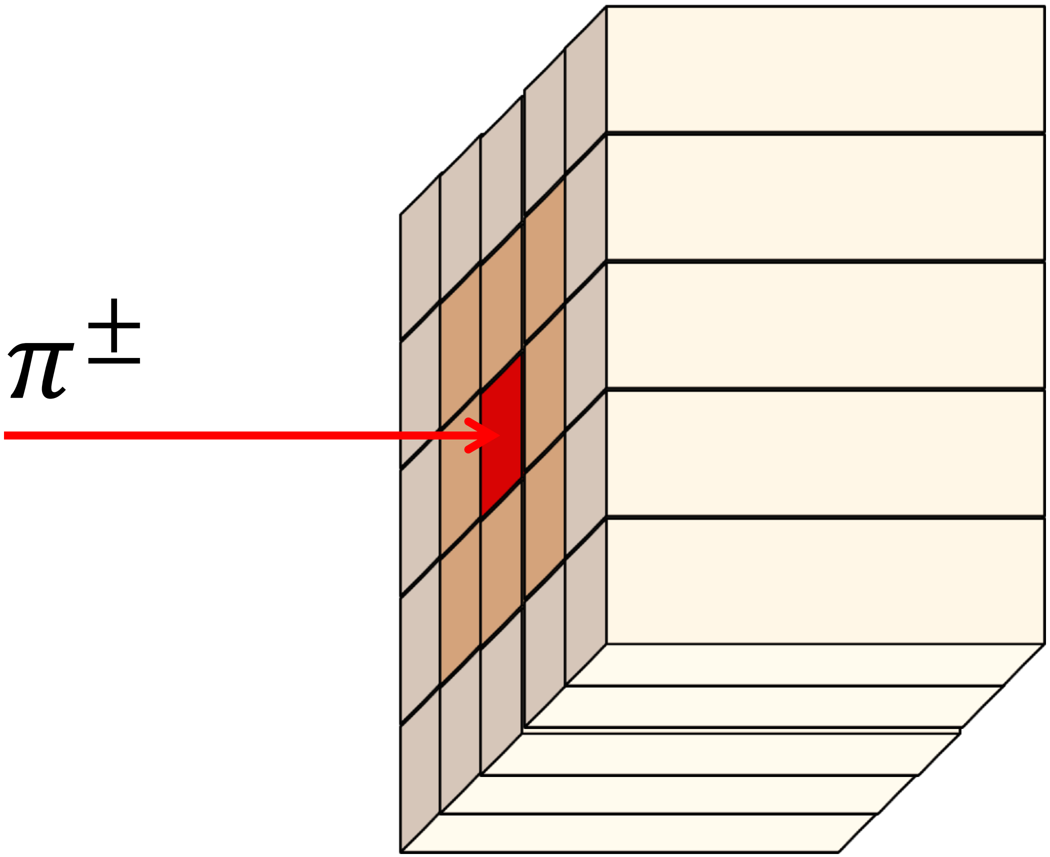


NO smearing

: all energy is deposited
in a single ECAL cell

Energy smearing into nearby ECAL cells

In reality



Smearing

: energy is smeared into nearby ECAL cells

- ✓ Sensitive to jet substructure variables

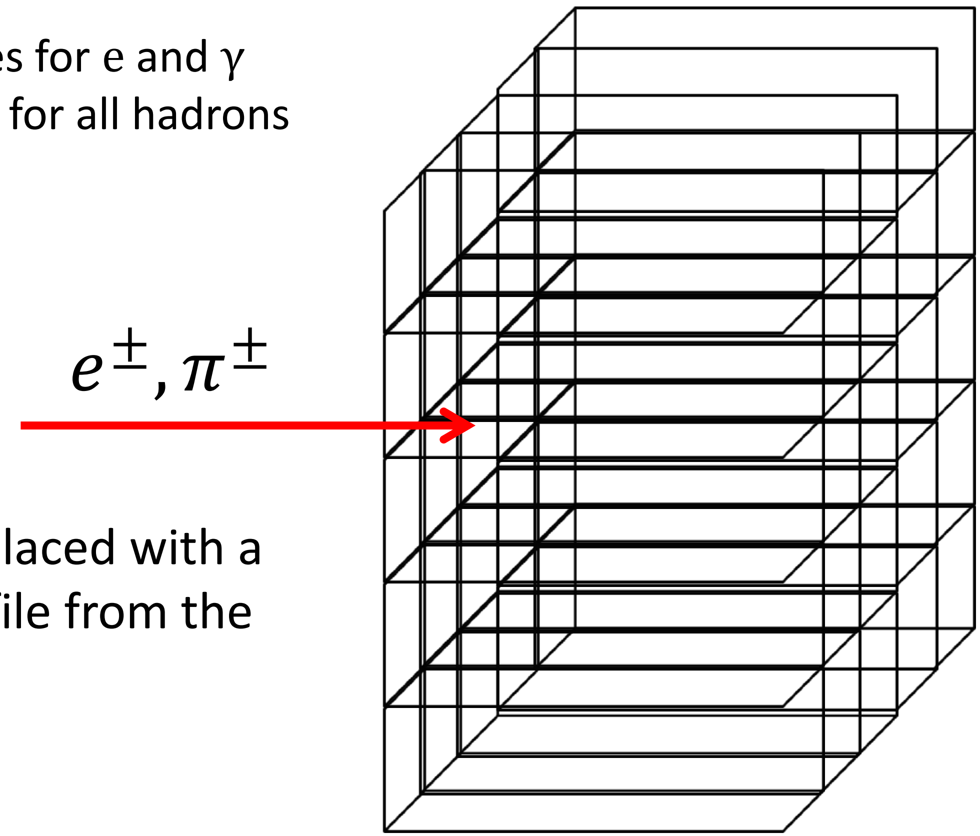
Smearing effect becomes extremely important in jet substructure analysis of the hyper-boosted heavy particles (e.g. top/H/Z/W)

We simulate ECAL smearing by GEANT

1. Prepare 9x9 ECAL cells with same dimension as CMS ECAL
2. Shoot single $e^{\pm}$, $\pi^{\pm}$ beams onto ECAL repeatedly
3. Build up a library of showering profiles for e, π beams

   ✓ e-induced showers as proxies for e and γ
   ✓ π-induced showers as proxy for all hadrons

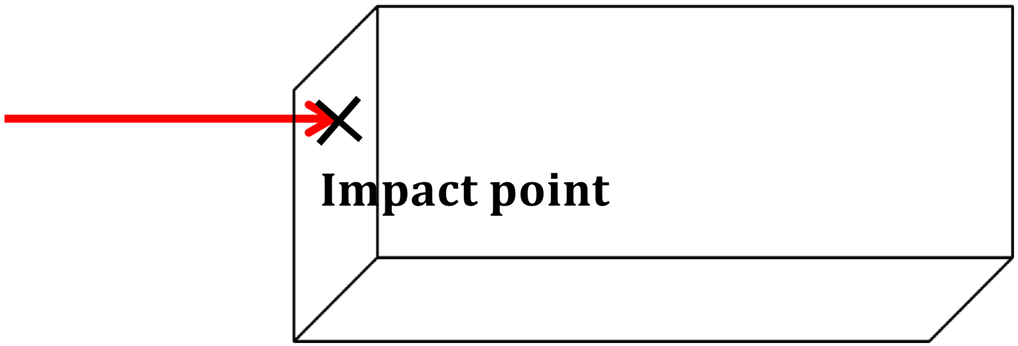
   Energy is fixed to be 100 GeV

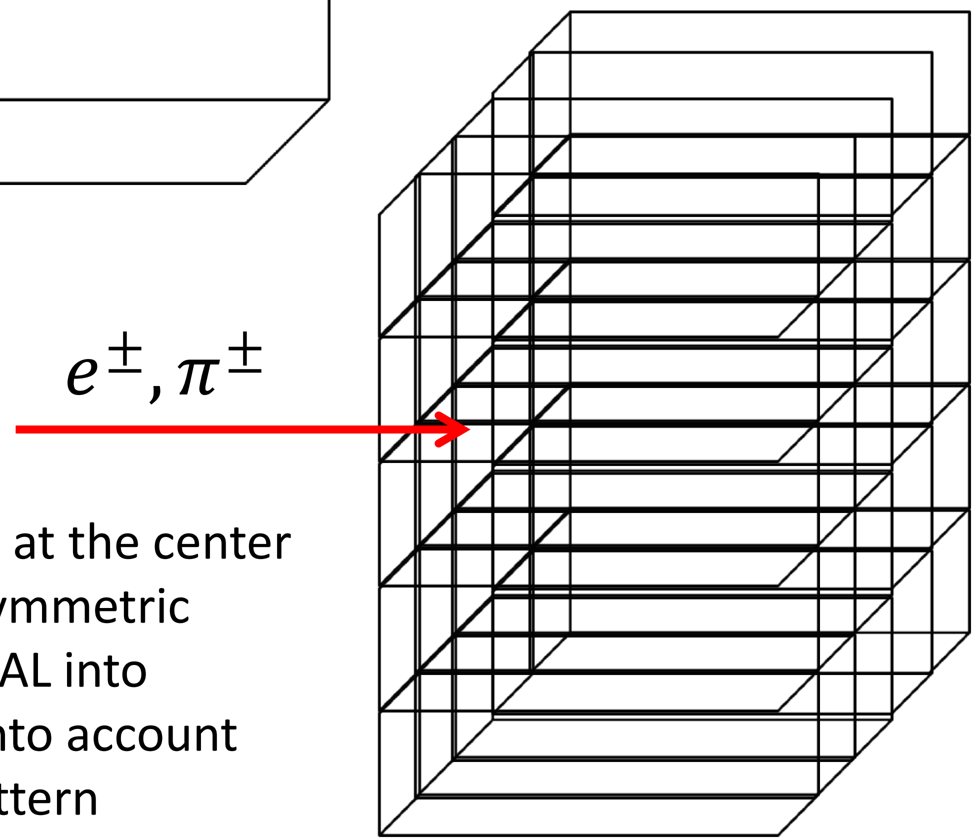


- ❑ Particle hitting a ECAL cell is replaced with a randomly chosen smearing profile from the library

*** Correlation between cells are automatically folded in**

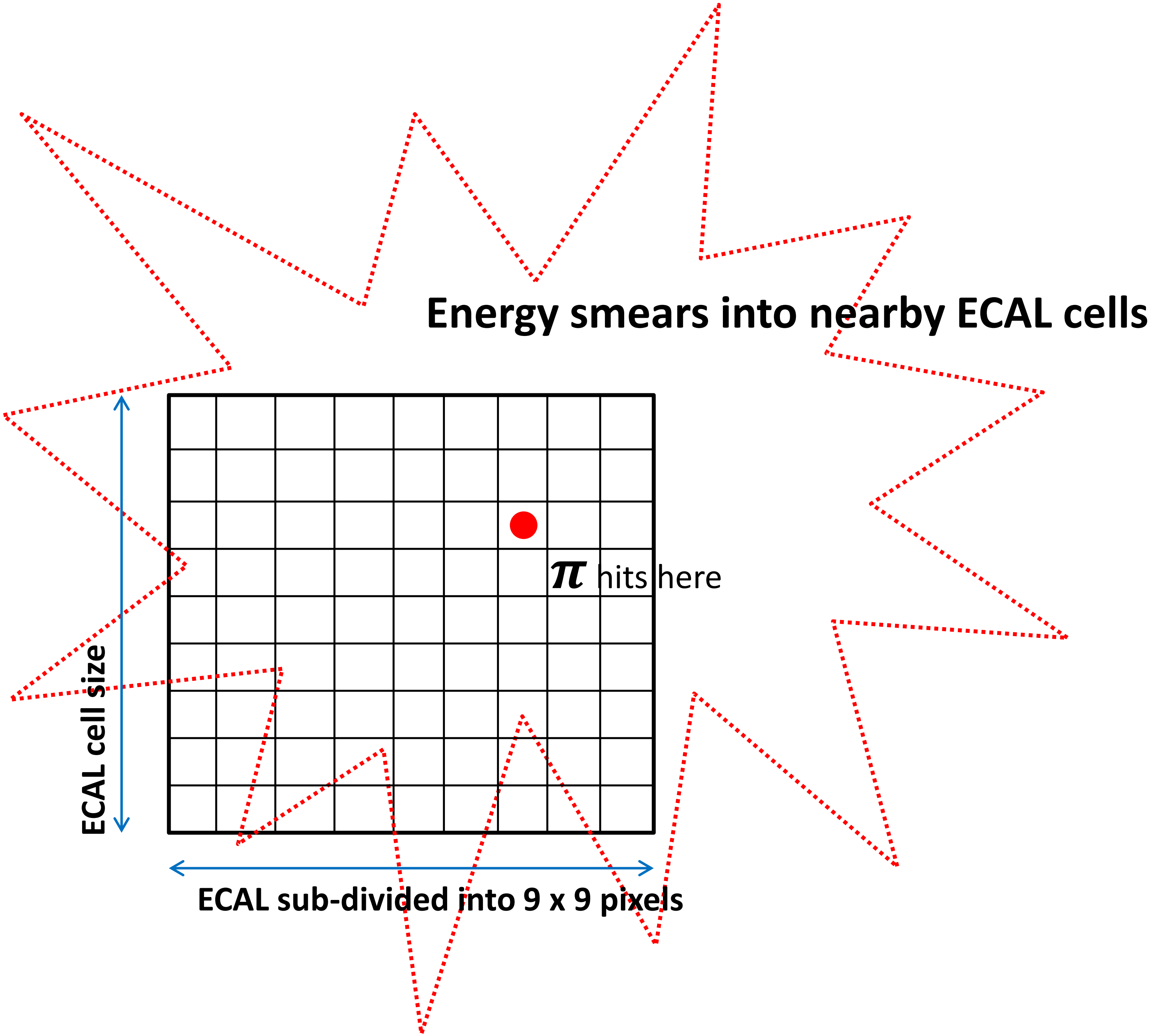
We simulate ECAL smearing by GEANT





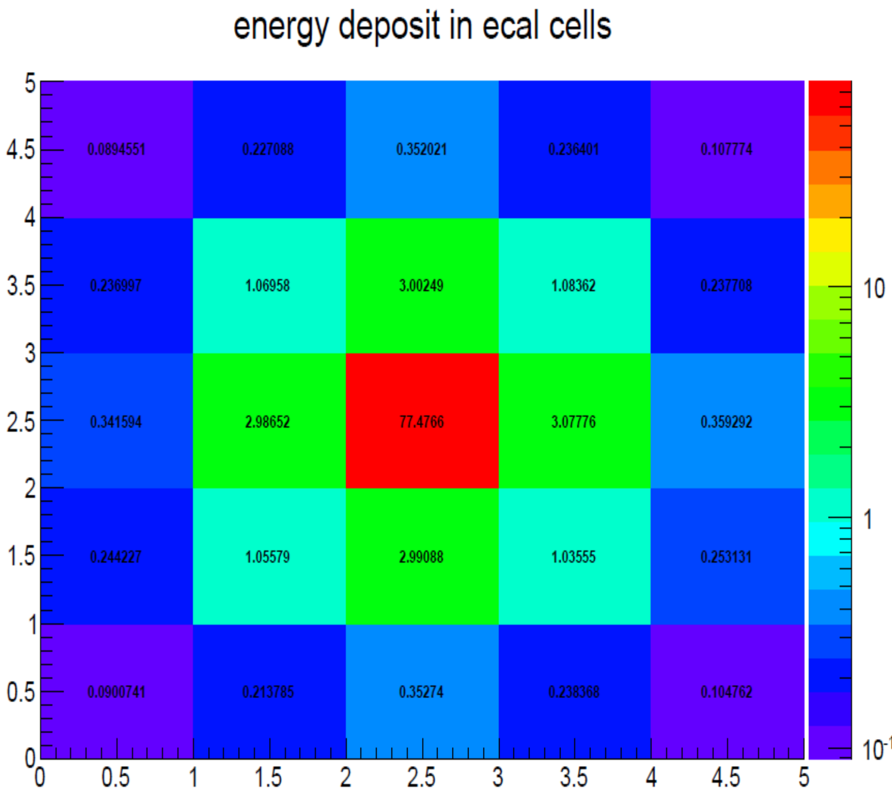
✓ Impact point is not always at the center of ECAL cell. It leads to asymmetric pattern. We sub-divide ECAL into 9x9=81 sub-cells to take into account asymmetric showering pattern

- We do not simulate asymmetric detector geometry, e.g. particle can hit a cell with an angle

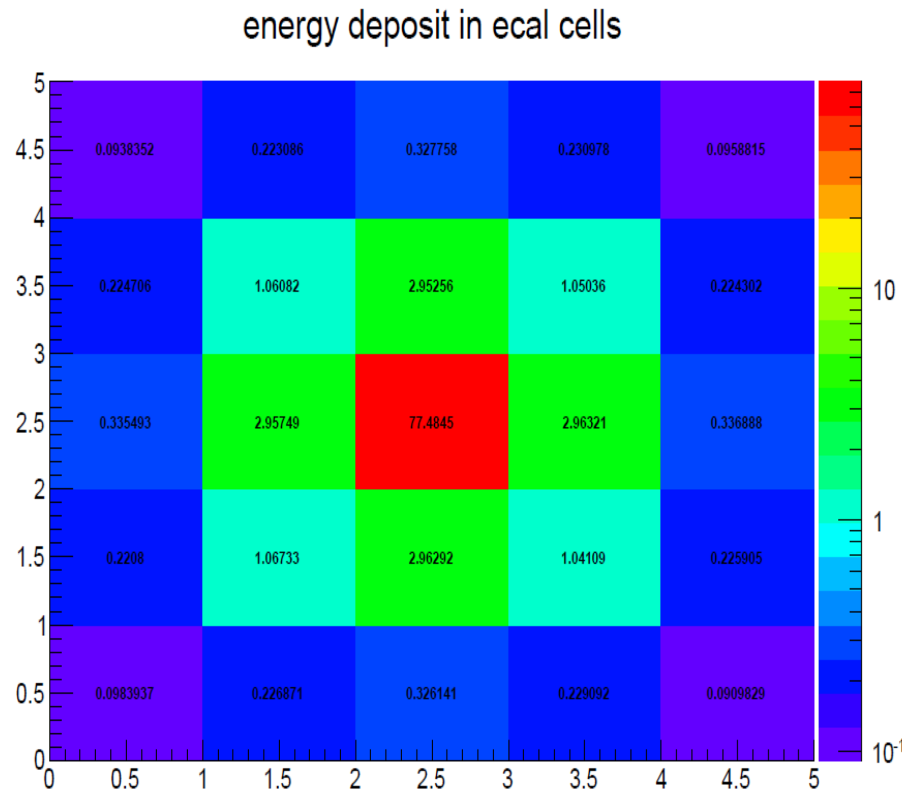

Energy smears into nearby ECAL cells
π hits here
ECAL cell size
ECAL sub-divided into 9 x 9 pixels

Electron-induced ECAL showering pattern by GEANT

10 GeV e^{-} beam



100 GeV e^{-} beam

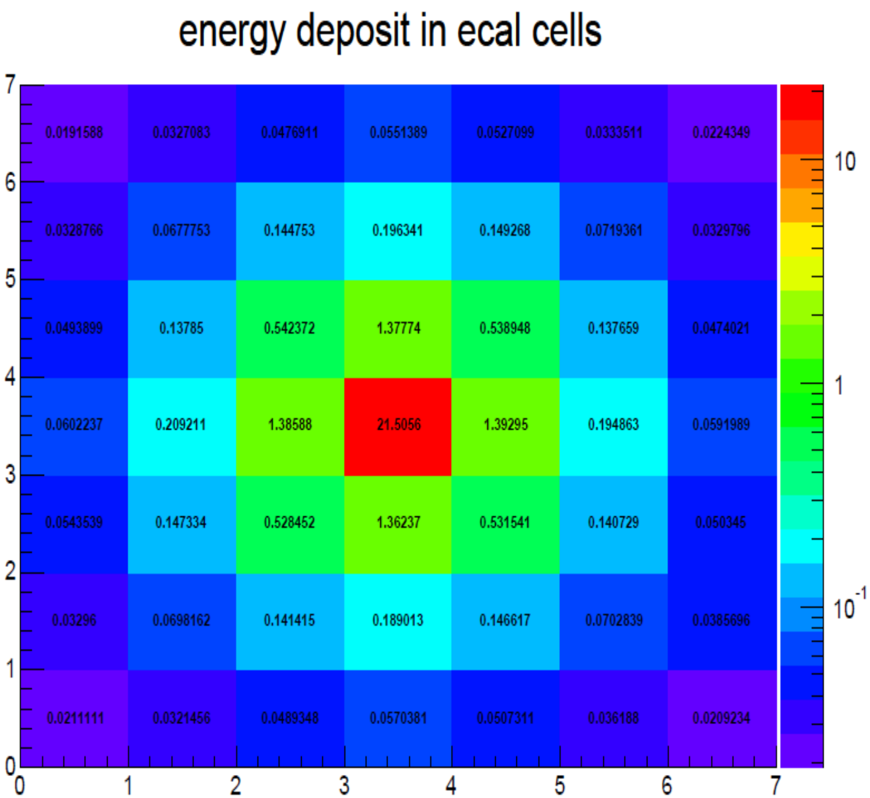


$E_{cell}/E_{incident\ electron}$, not w.r.t $E_{total\ deposit}$

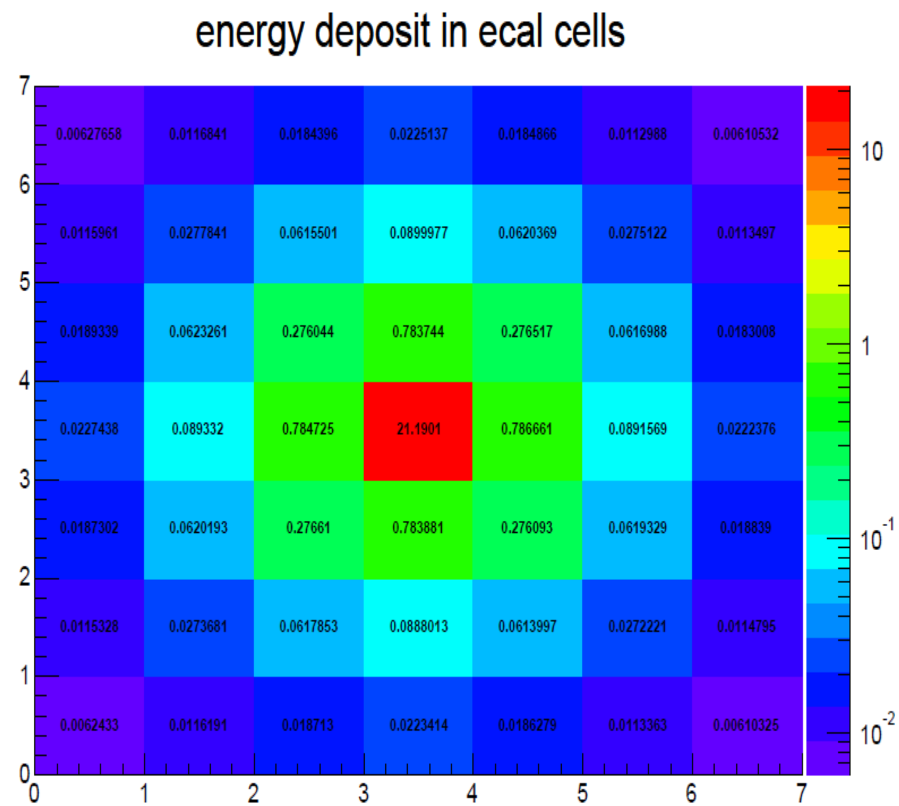
- Nearly pT-independent. It justifies our proxies simulated at 100GeV

Pion-induced ECAL showering pattern by GEANT

100 GeV $\pi^{\pm}$ beam



3 TeV $\pi^{\pm}$ beam

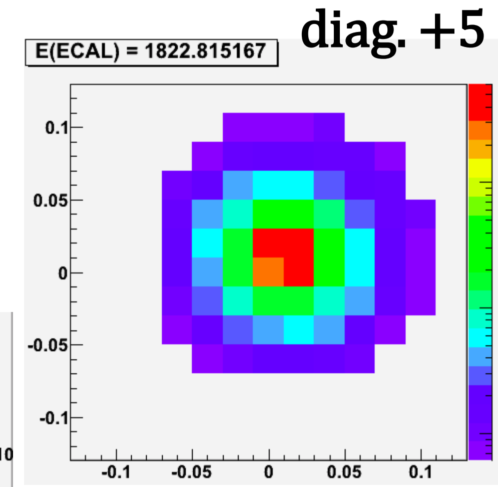


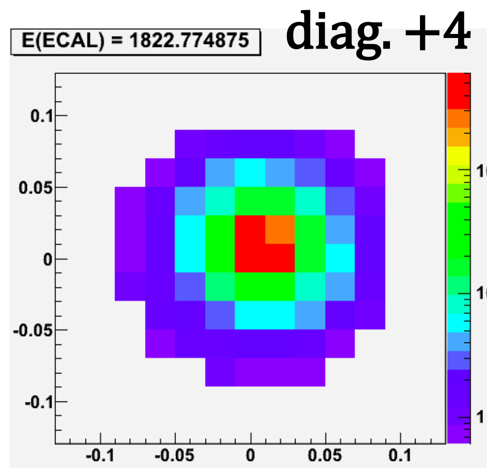
$E_{cell}/E_{incident\ pion}$, not w.r.t $E_{total\ deposit}$

- Nearly pT-independent. It justifies our proxies simulated at 100GeV

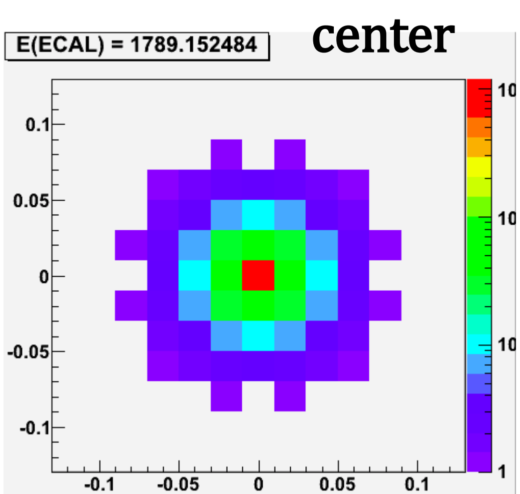
Pion events in our library approach

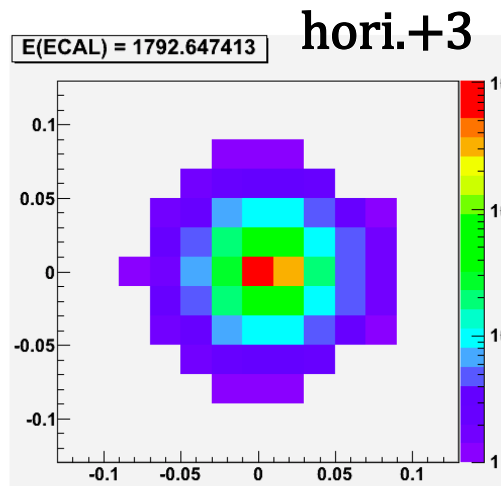
Asymmetric smearing pattern over 9 × 9 ECAL cells caused by random impact point away from the center

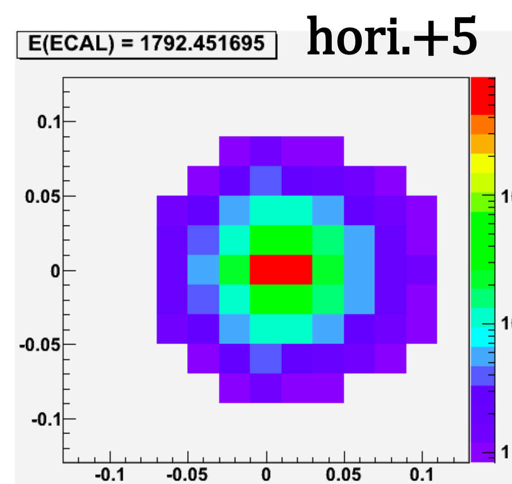




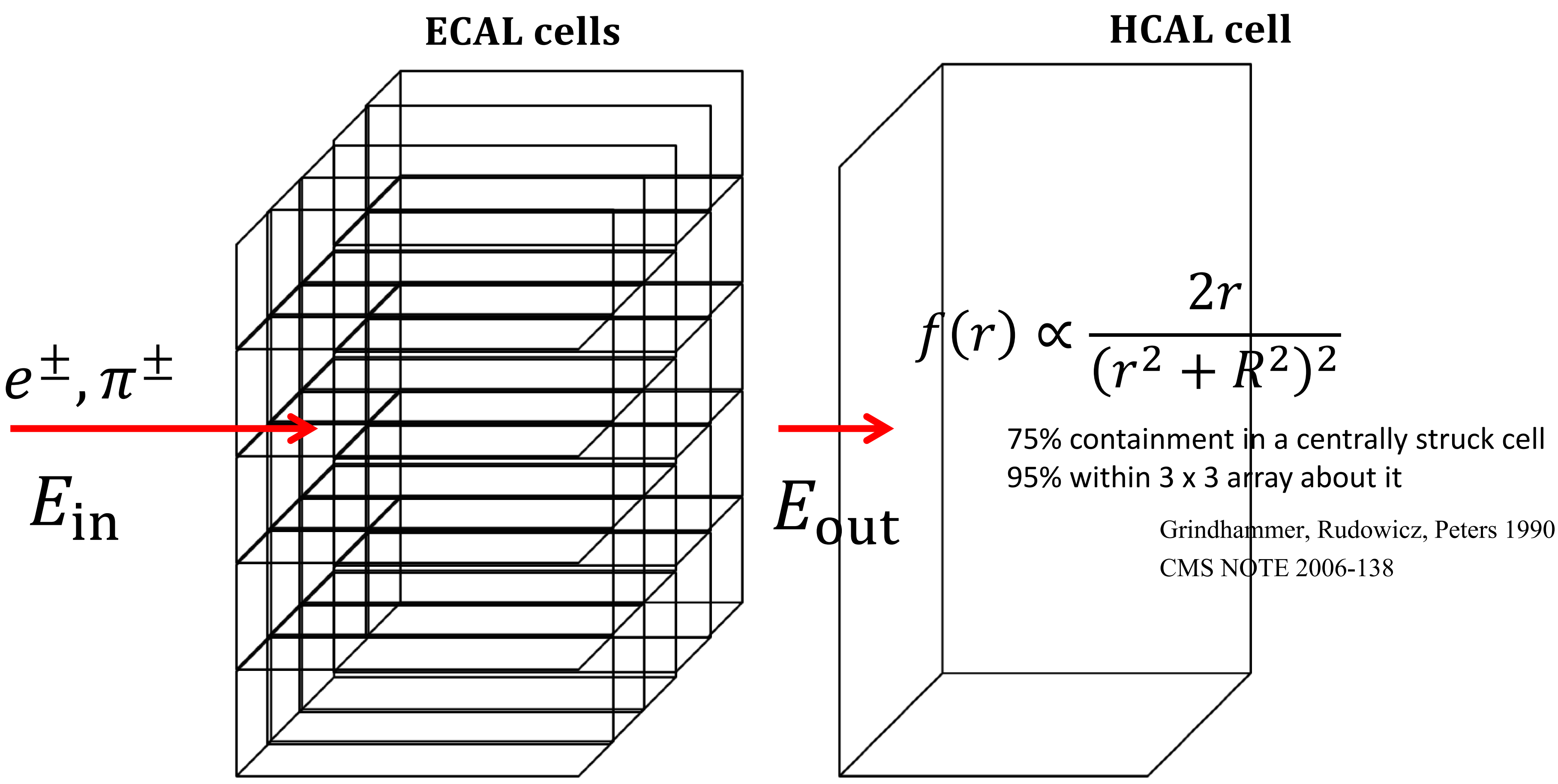
Accumulated plot of 100 pions







Profile ansatz for HCAL



$$f(r) \propto \frac{2r}{(r^2 + R^2)^2}$$

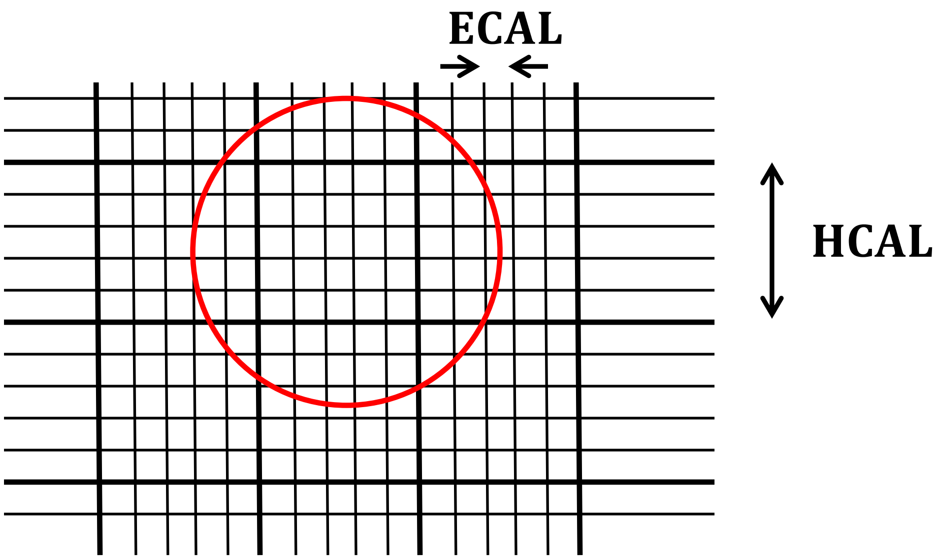
Replace all particles flowing out the back of an ECAL cell with a continuous angular energy distribution according to the above ansatz

Spurious structure due to smearing

Smearing into nearby cells can introduce spurious structure when a rescaling is done within each HCAL cell

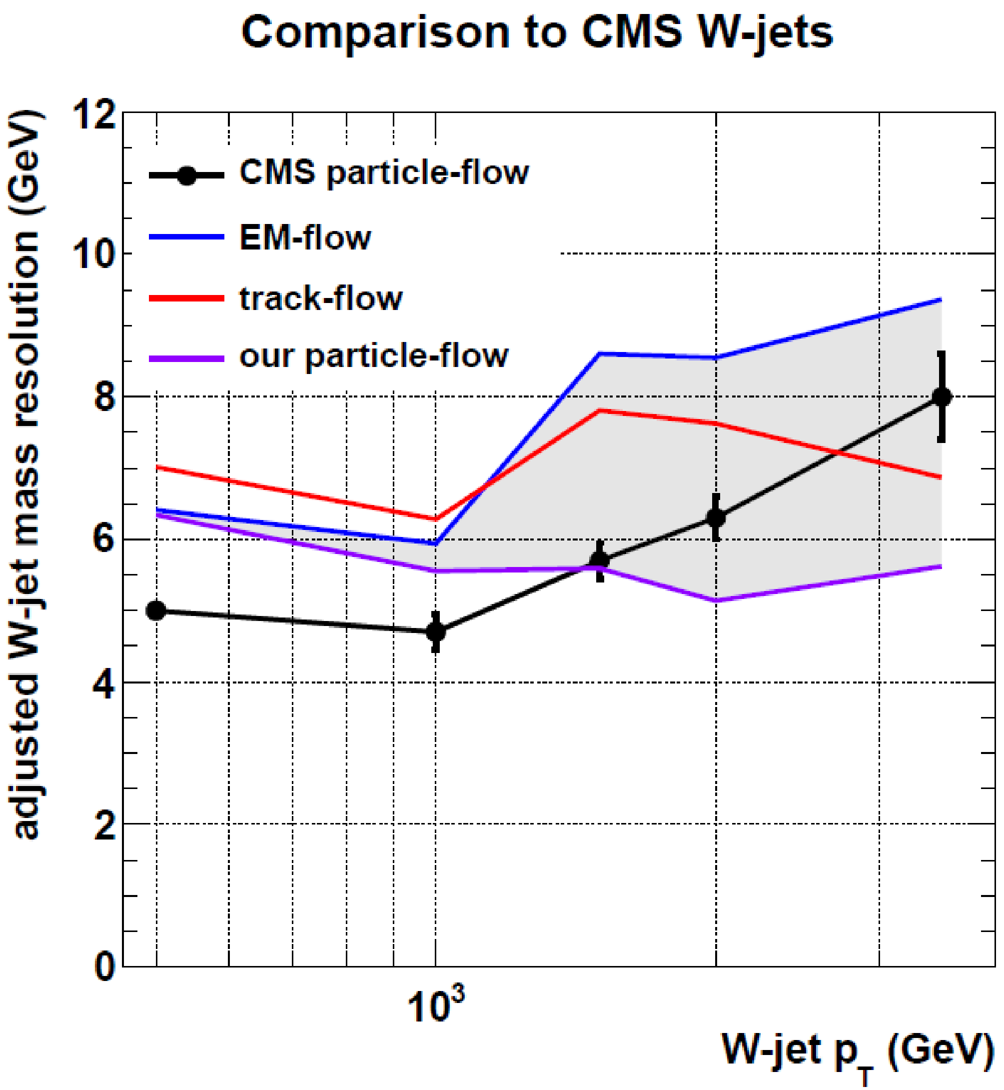
Mini-jet clustering

- ✓ deals with HCAL energy spreading:
  e.g. in EM-flow, the entire collection of ECAL and HCAL cells are clustered into mini-jets with the anti-k_T algorithm with the size comparable to the HCAL size. Rescaling is carried out within each mini-jet



Validation of our approach against CMS high pT W-jet

CMS PAS JME-14-002



Three benchmark scenarios

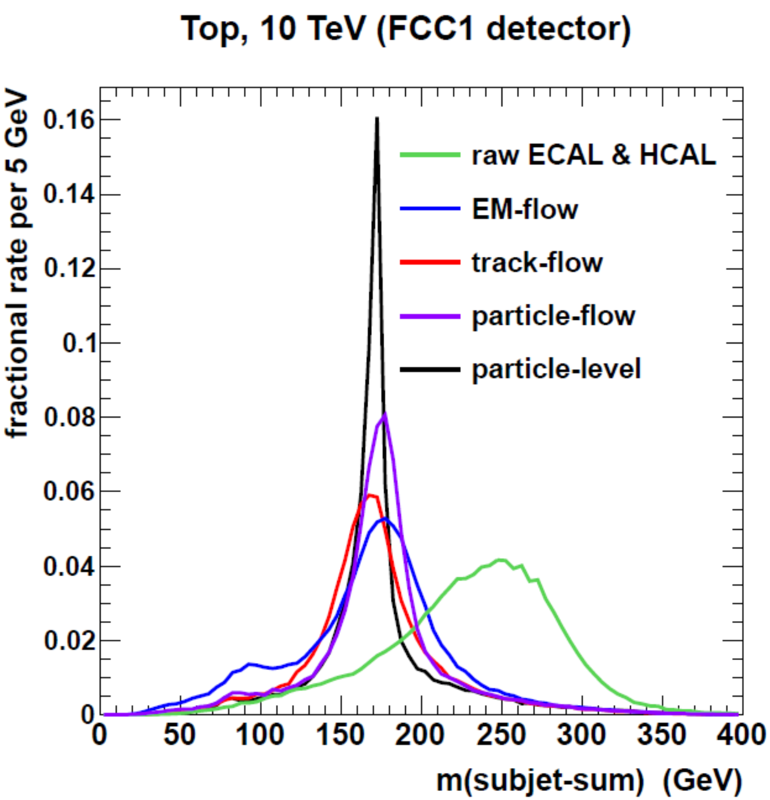
Model	Tracking: two extremes	ECAL material	ECAL cell	HCAL cell
LHC		CMS-type ($PbWO_4$)	0.02×0.02	0.1×0.1
FCC1	Perfect/absent	$PbWO_4$ (Lead tungstate)	0.01×0.01	0.05×0.05
FCC2	Perfect/absent	Pure W (Tungsten)	0.005×0.005	0.05×0.05

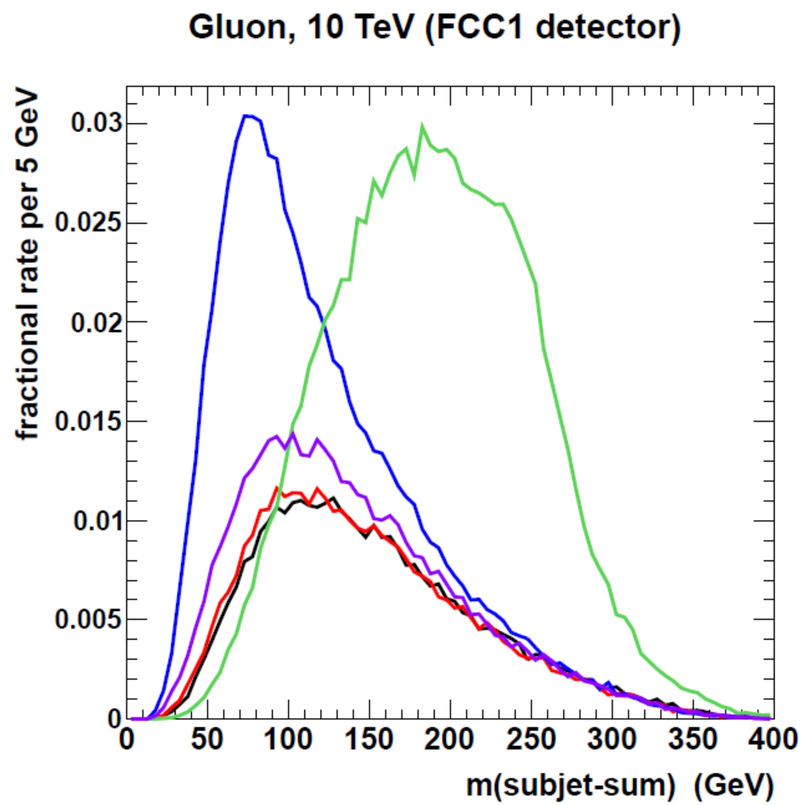
We will see how these detector models perform in three benchmark LHC/FCC detectors

- ❑ **Raw ECAL & HCAL**
- ❑ **EM-flow**
- ❑ **Track-flow**
- ❑ **Particle-flow**
- ❑ **Particle-level**

- Effective Moliere radius of pure W is bigger than what is assumed. Consider Pure W as a place-holder for any new material with a half-sized effective Moliere radius

Filtered top-jet mass of 10TeV top/gluon at FCC1

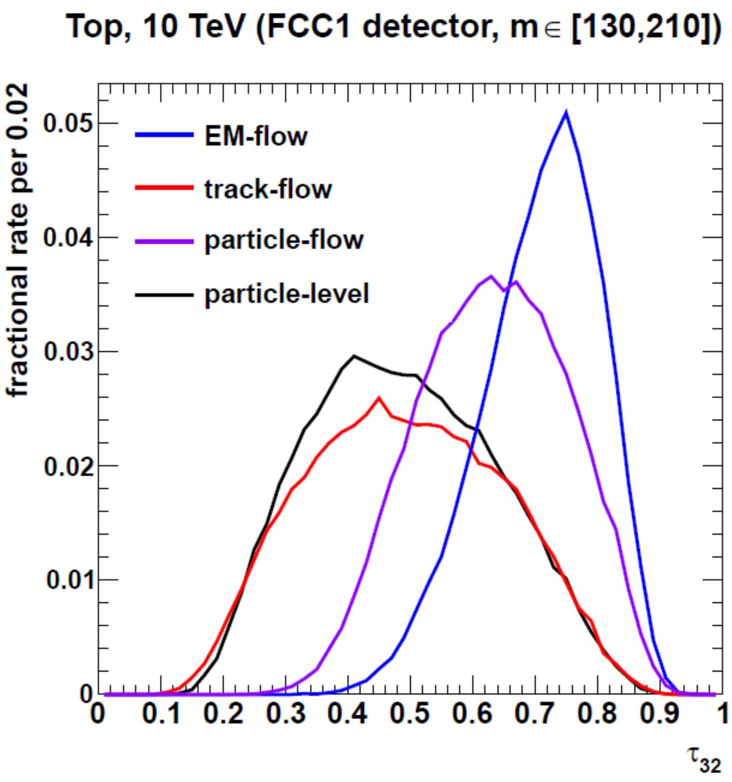


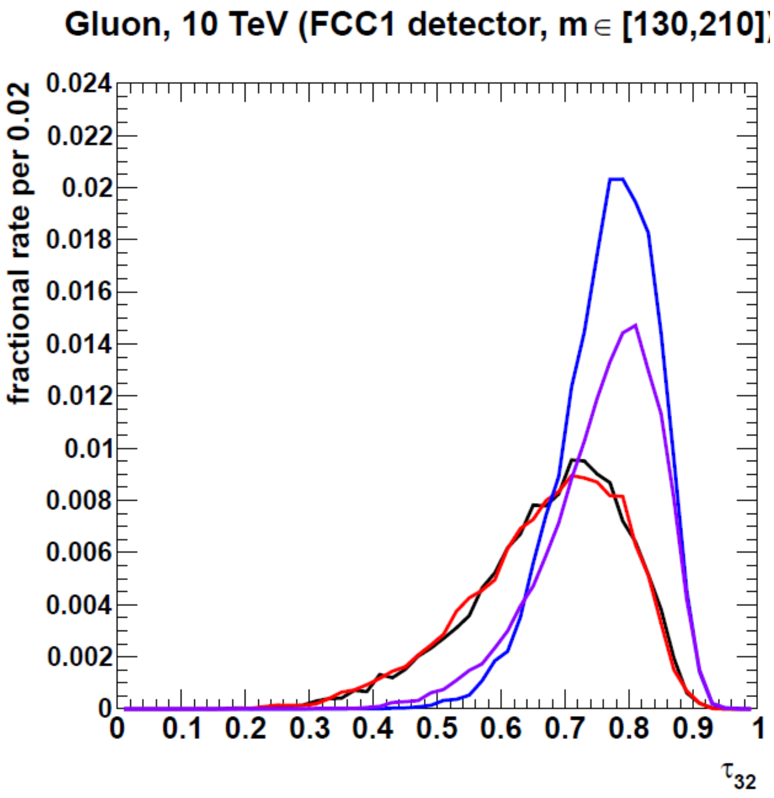


This situation is equivalent to 5TeV top/gluon at the LHC

- pile-up and magnetic field are not included in this study

τ_{32} of 10TeV top/gluon at FCC1

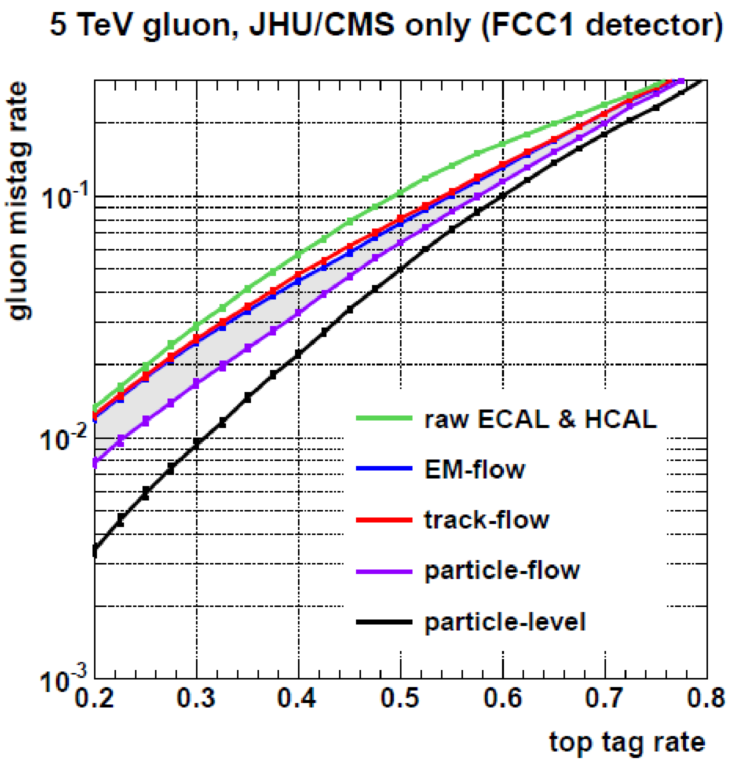


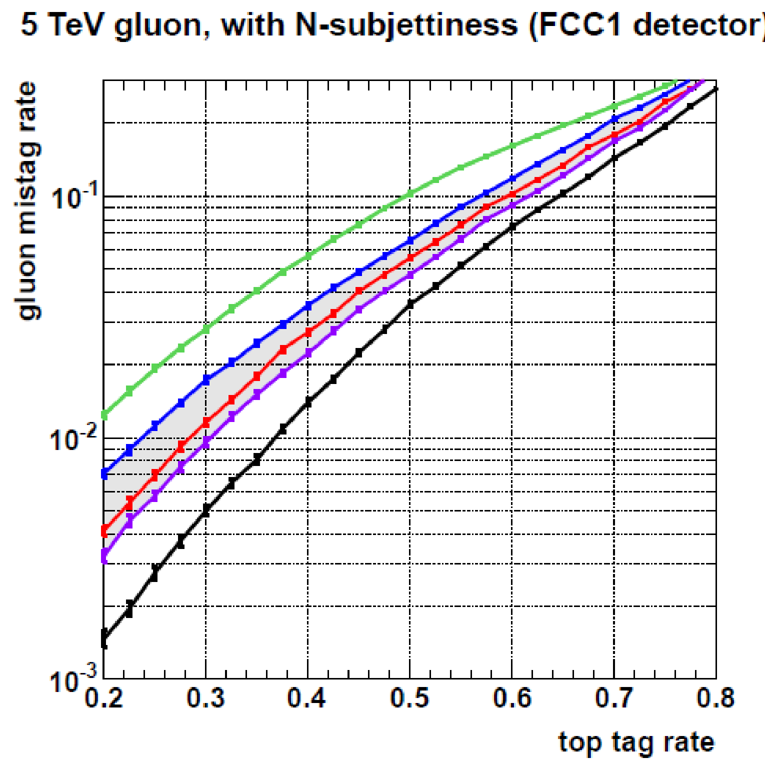


Note that N-subjettiness is doing great whenever tracks are available

- Perfect tracking efficiency is assumed in track-flow
- τ_{32} seems to probe a property within JHU/CMS subjets, rather than in-between them

5TeV top/gluon discrimination at FCC1

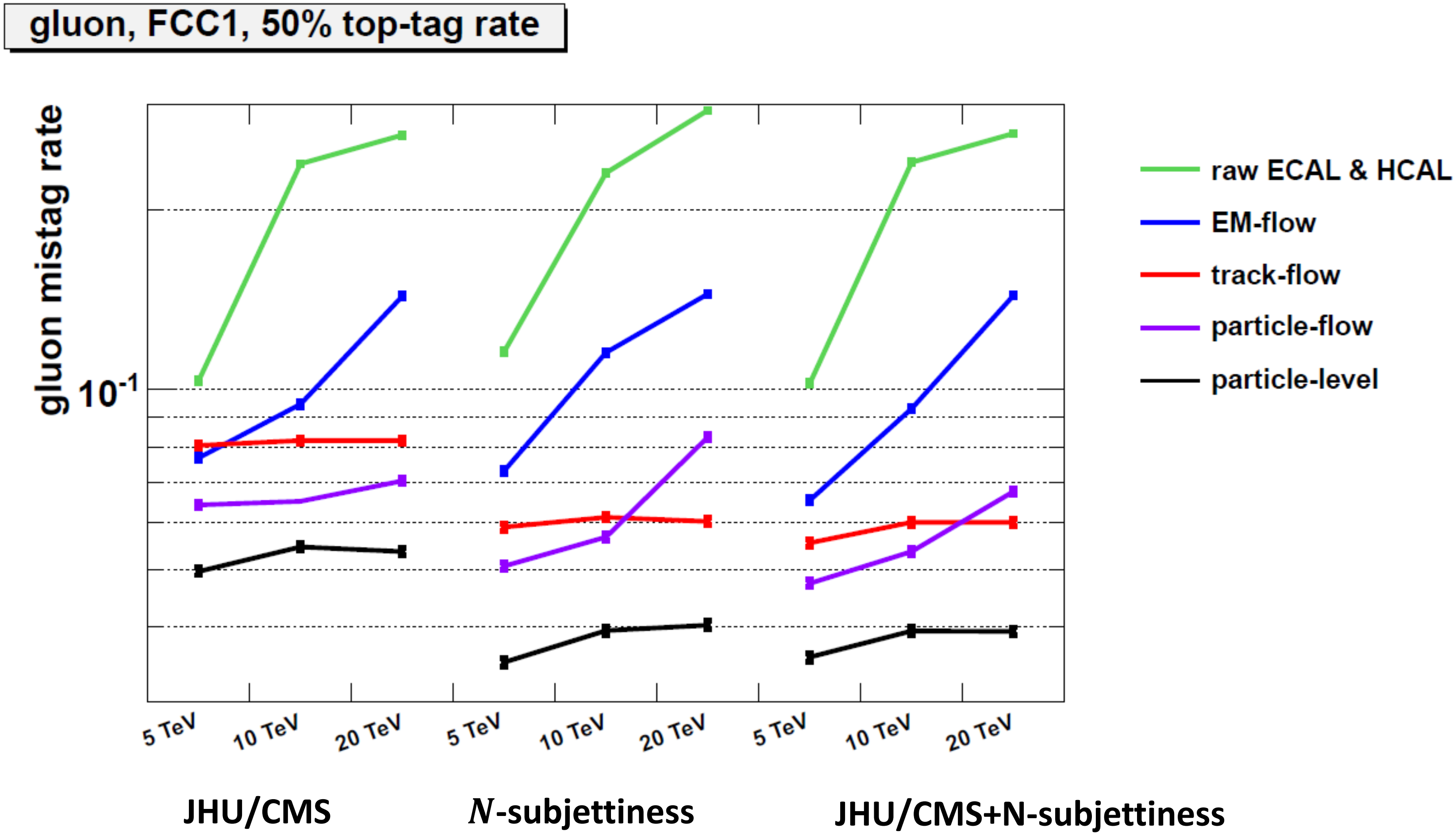




✓ equivalent to 2.5 TeV top/gluon-jets at the LHC

- Particle-flow is universally the best option
- Track-flow seems to work better with N-subjettiness, EM-flow is less effective at capitalizing on N-subjettiness

- The gap b/w particle-flow and particle-level is driven by that m_t and m_W invariant mass features become less sharply-peaked for the top jets

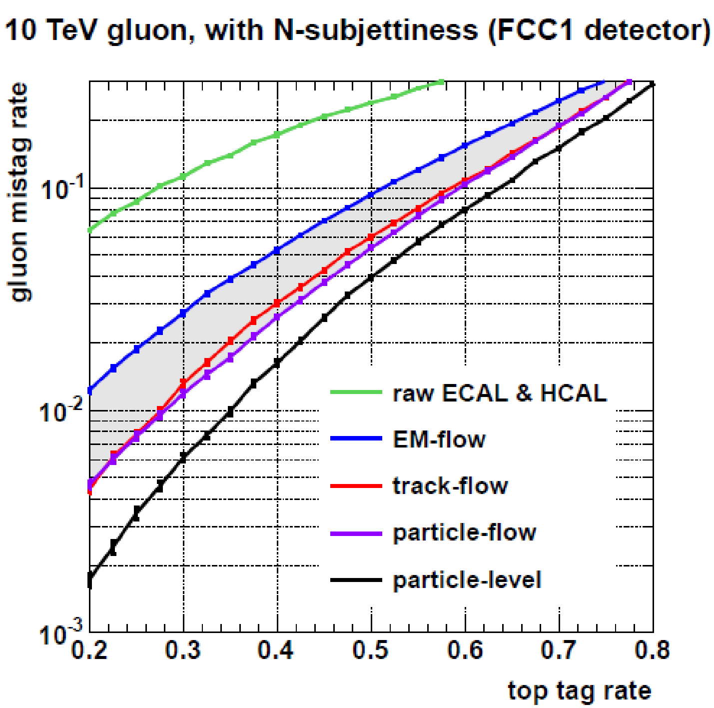


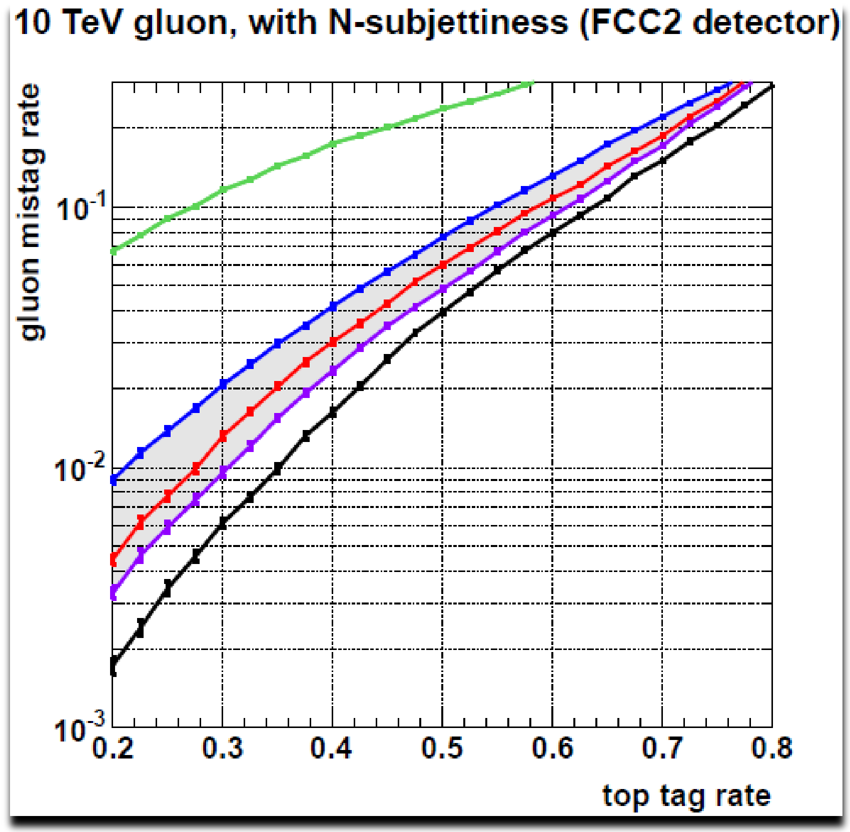
- JHU/CMS tagger never fully competitive with N-subjettiness (except for EM-flow at 10TeV)
- Combined tagger is universally better

✓ FCC2 brings EM-flow, particle-flow to the similar level of half-p_T jets at FCC1

10TeV top/gluon discrimination at ``FCC1 → FCC2”

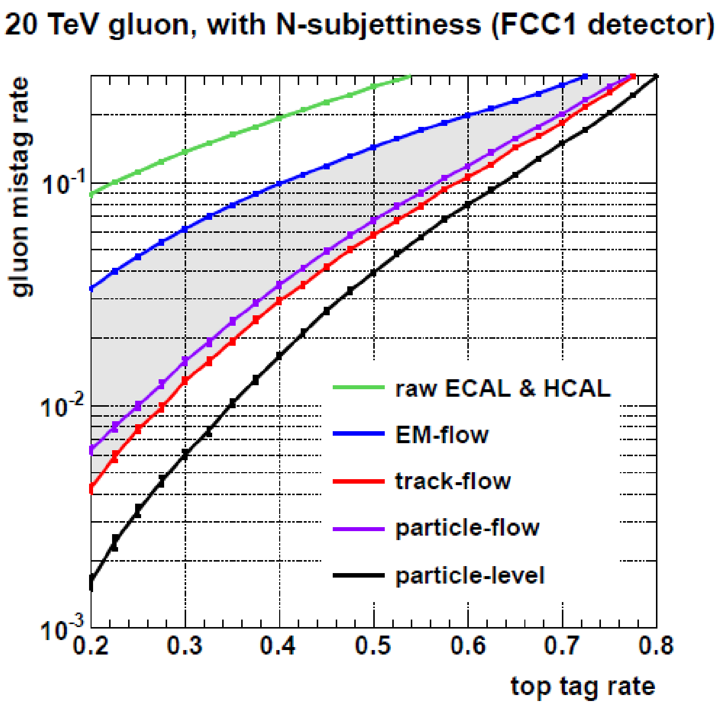
ECAL 2x, HCAL 1x

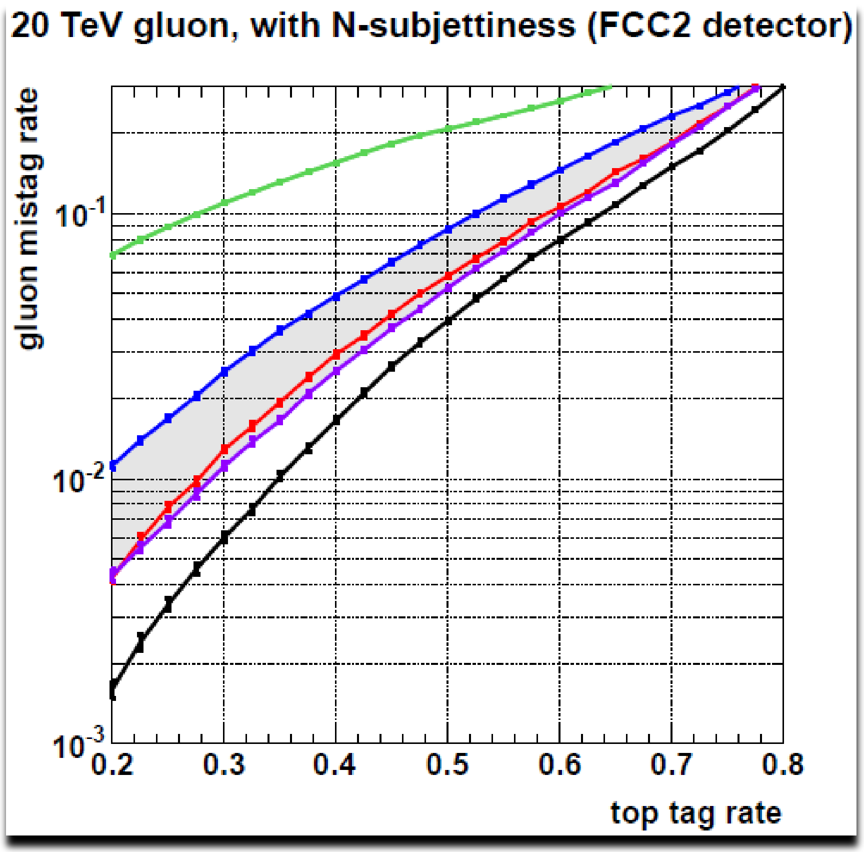




- ✓ Performance of EM-flow, particle-flow get restored to that at 5 TeV when going to FCC2 from FCC1

20TeV top/gluon discrimination at ``FCC1 → FCC2”

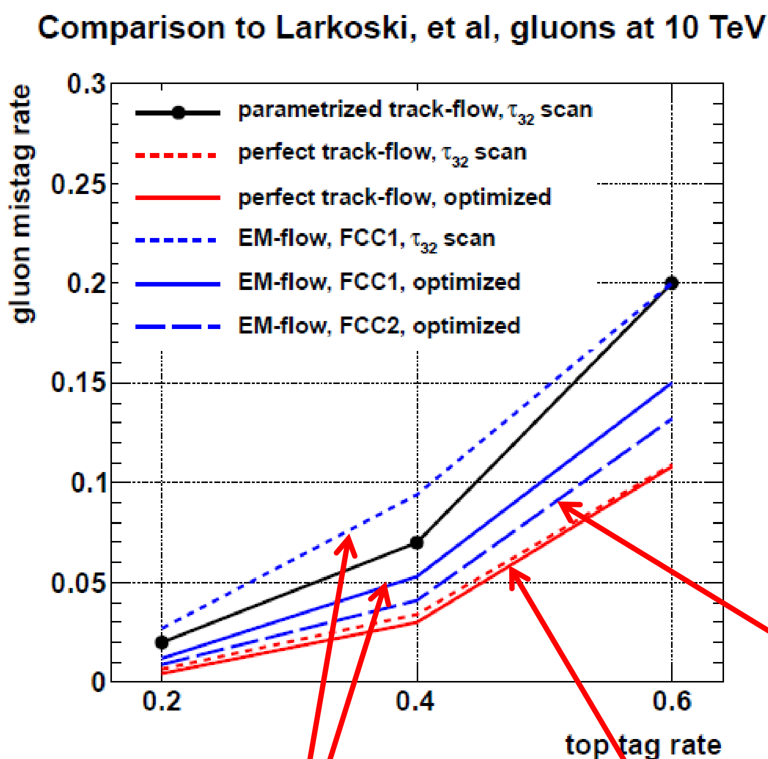


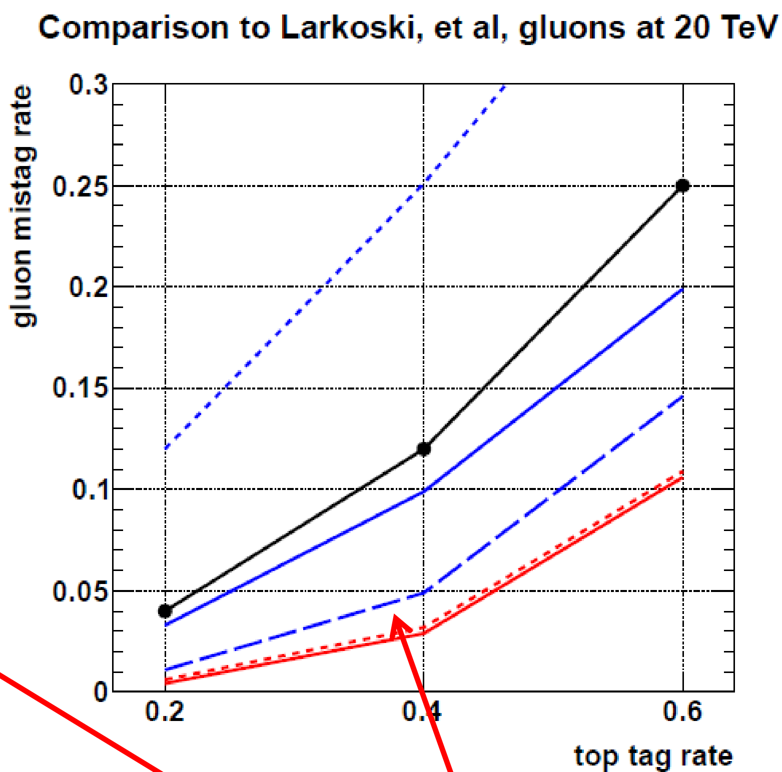


✓ FCC2 at 20TeV looks similar to FCC1 at 10TeV

Comparison to an existing study

Larkoski, Maltoni,Selvaggi 2015





When tracks are not available, JHU/CMS tagger does most job

When tracks are available, τ_{32} does most job

Stable. EM-flow can cover up to 20 TeV at FCC2

Strong Magnets at FCC

✓ Beneficial to high-p_T physics. It hurts low-p_T physics

	CMS: 4T, 1.5m	FCC: 6T, 6m
$p_{T\ crit} = 0.15 \times \left(\frac{\mathrm{B}}{\mathrm{T}}\right) \times \left(\frac{\mathrm{r}_{cal}}{m}\right)$	~ 0.9 GeV	~ 5.4 GeV

- This implies that $O(100\ \mathrm{GeV})$ process such as Higgs physics becomes low-p_T physics at 100 TeV!

  E.g. $H \to b\bar{b}$ with low p_T will be significantly under-reconstructed due to lost tracks (We need to make sure that we are capable of restoring the lost tracks back to our jets via track reconstruction, e.g. particle-flow)

In a situation that strong magnetic field becomes problematic, it hurts high-p_T tracking efficiency, but

EM-flow is insensitive to this issue

To conclude

- The performance of our optimization of JHU TopTagger combined with N-subjettiness

  1. Quark- and gluon-jets can be simultaneously optimized within JHU TopTagger
  2. Adding N-subjettiness to e.g. JHU TopTagger, can make $O(1)$ improvement of top/gluon discrimination
  3. N-subjettiness is effective when tracks are available
  4. JHU is more robust than N-subjettiness under more pessimistic detector assumptions

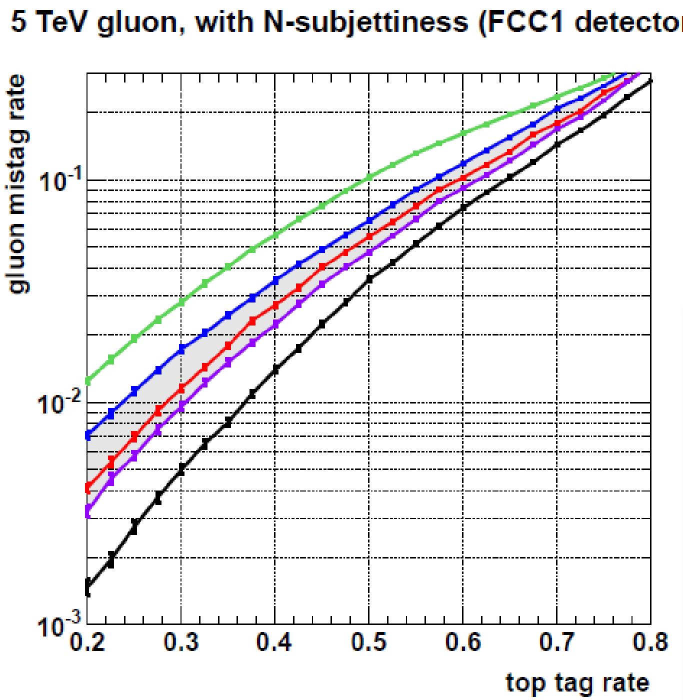
- EM-flow looks very promising. It can solely cover up to 20TeV tops assuming FCC2 configuration (ECAL 4x, HCAL 2x)

  1. Trackers become crucial to tag tops beyond it
  2. Unless the FCC detectors are constructed with near-perfect trackers, some additional investment in ECAL granularity would be beneficial

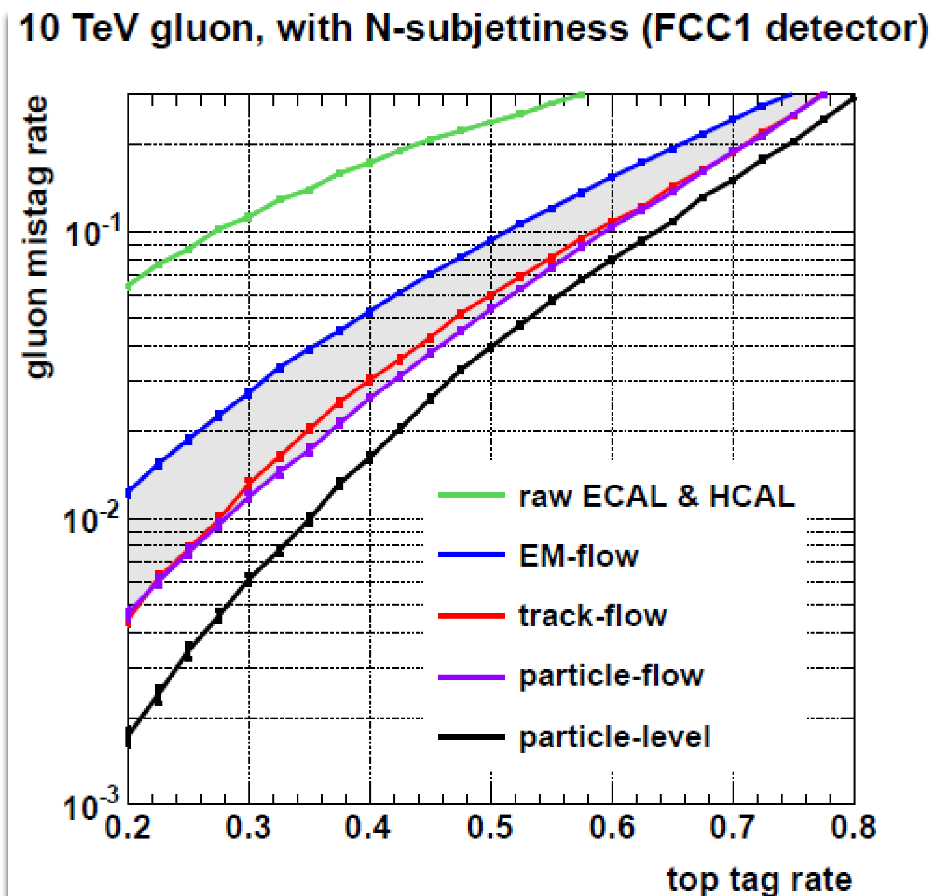
- An issue on W-strahlung/FSR etc will be discussed in our paper

Extra Slides

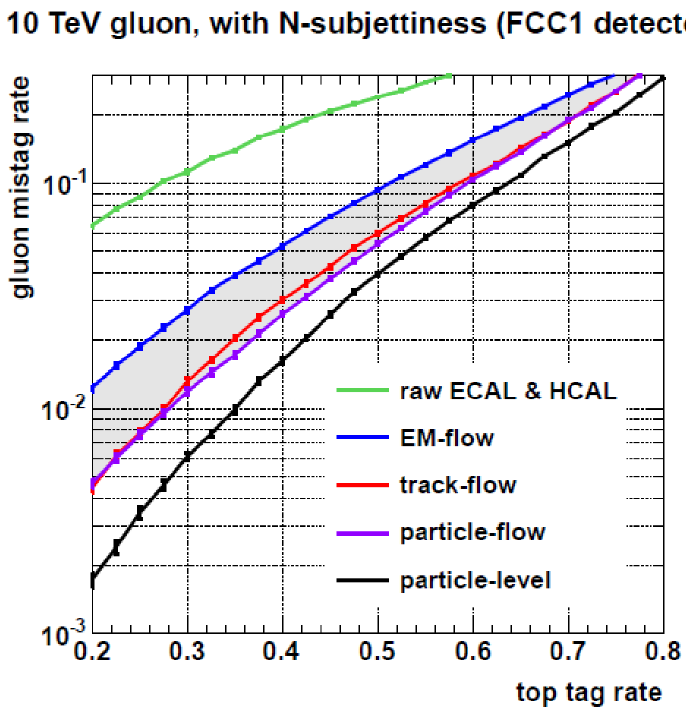
``5TeV→10TeV'' top/gluon discrimination at FCC1



- Track-flow, particle-level stay same (should be scale-invariant) w.r.t. 5 TeV as we assume perfect tracking
- Raw ECAL & HCAL gets weakened a lot
- EM-flow, particle-flow also get weakened (less pronounced)



``10TeV→20TeV'' top/gluon discrimination at FCC1



- Track-flow, particle-level stay same as 5TeV to 10TeV
- The performance of EM-flow is limited
- Our simple Particle-flow performs worse than track-flow