

Jet Reconstruction Performance & Distinguish between Multi-jet Events



Pei-Zhu Lai



NCU (Taiwan)

On the behalf of the CEPC Collaboration

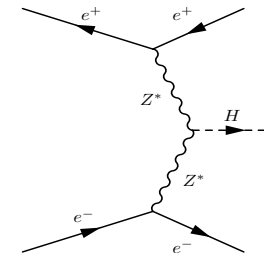
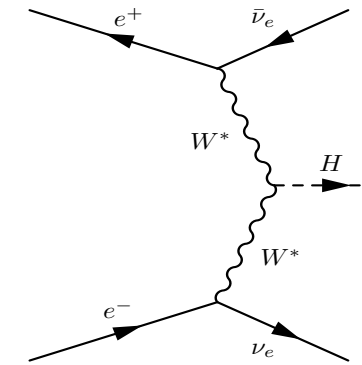
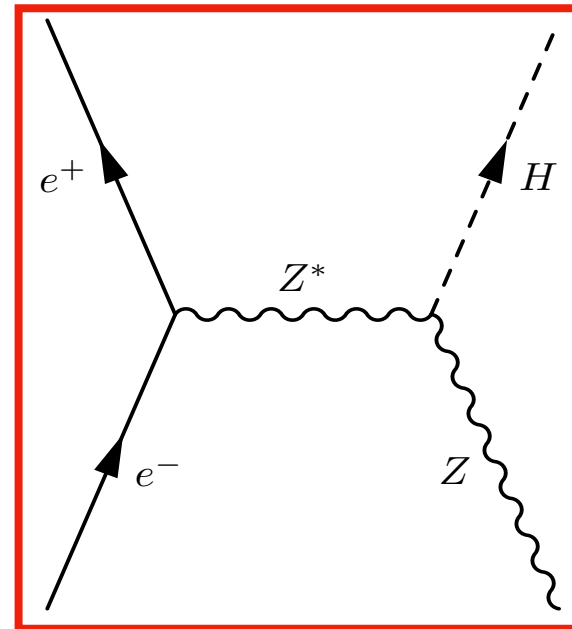
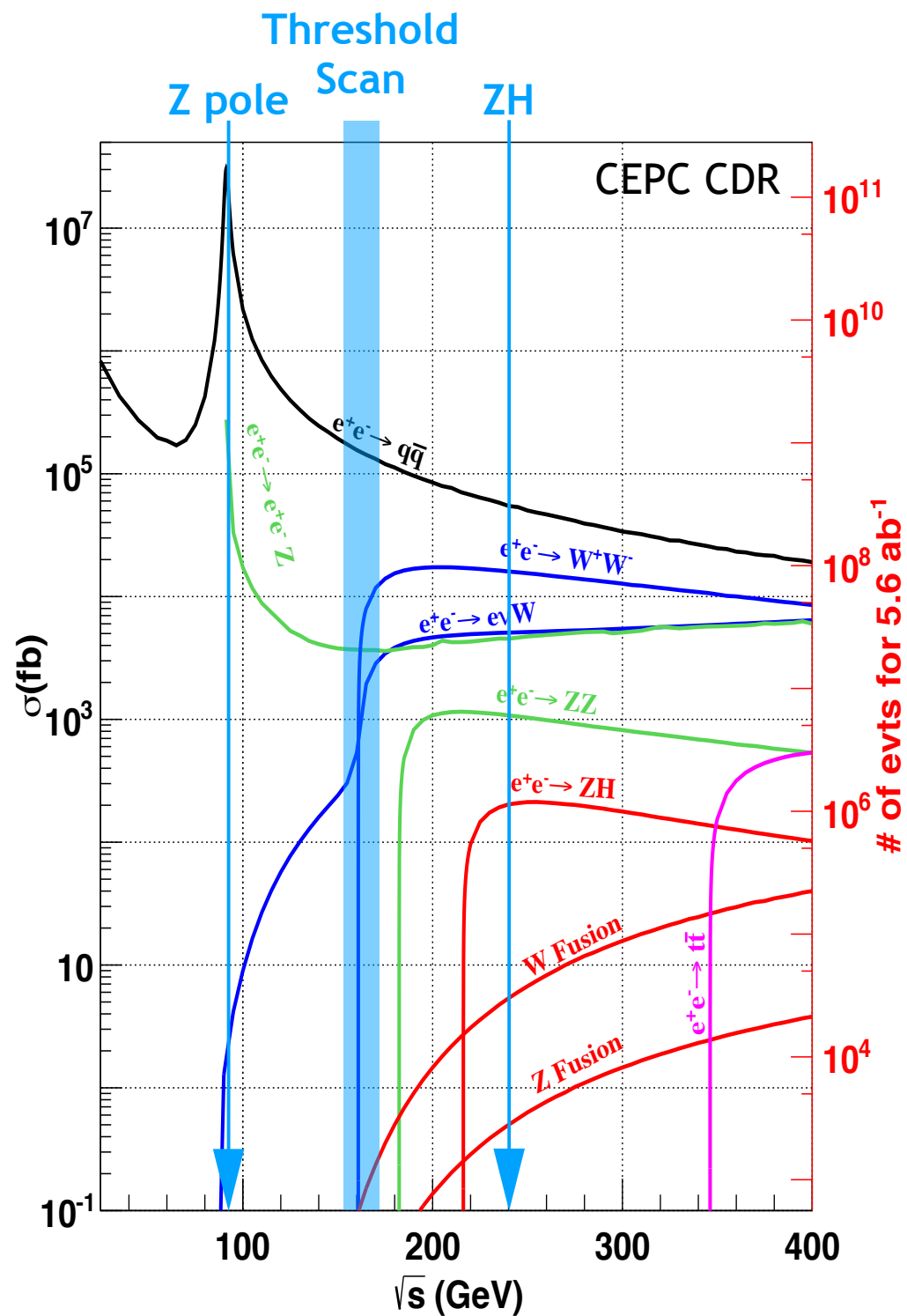
(pei-zhu.lai@cern.ch)

LCWS2019, Sendai International Center, Sendai Japan

Oct 28 - Nov 01, 2019

- Higgs production at CEPC
- ZH decay mode
- Jets at the Higgs Signal
- Jet performance in different physic benchmark
- Summary

Higgs Production at CEPC

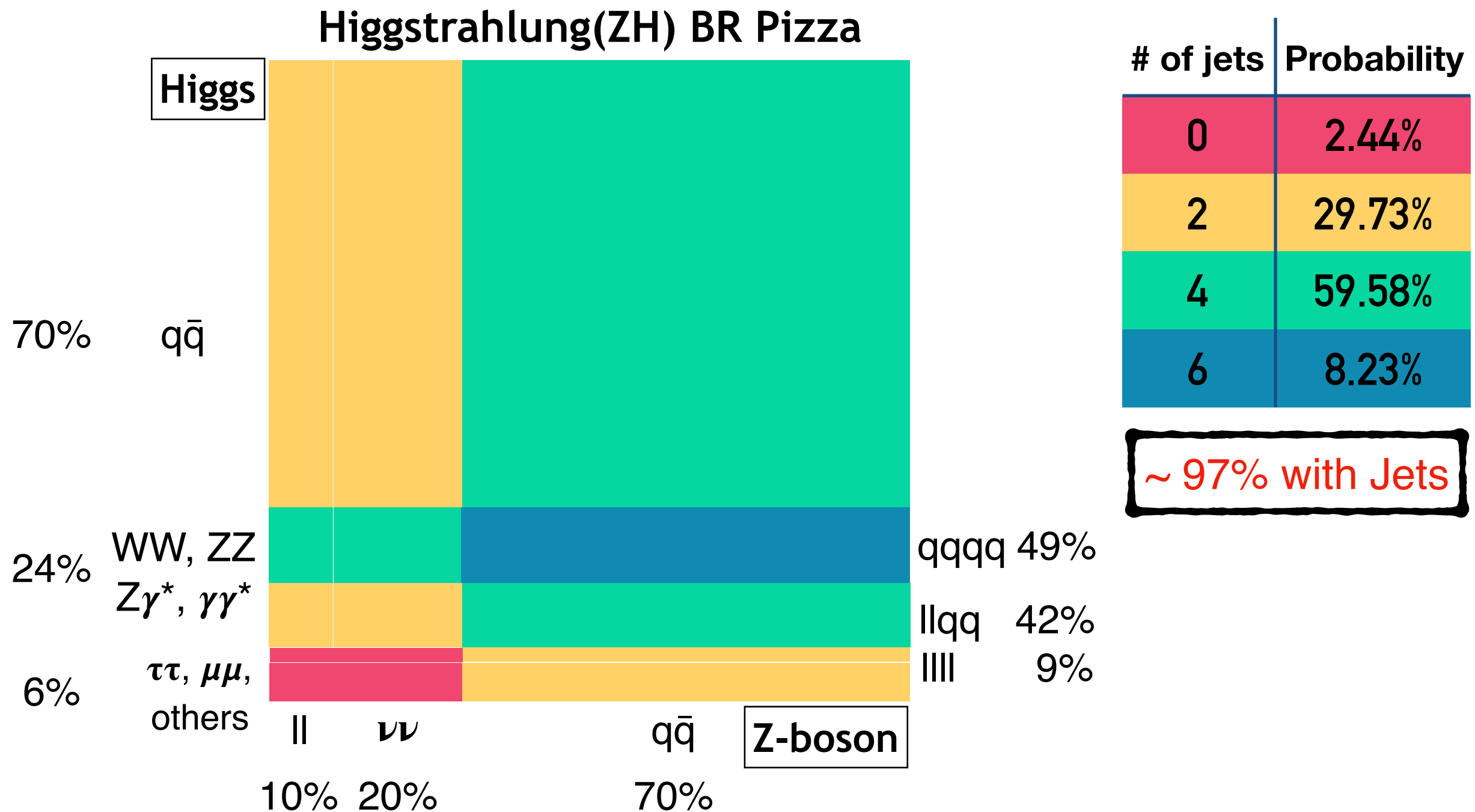


Process	Cross section(fb)	Events in 5.6 ab ⁻¹
$e^+e^- \rightarrow ZH$	196.2	1.10×10^6
$e^+e^- \rightarrow \nu_e \bar{\nu}_e H$	6.19	3.47×10^4
$e^+e^- \rightarrow e^+e^- H$	0.28	1.57×10^3
Total	203.7	1.14×10^6

$$S : B = 1 : (100 \sim 1000)$$

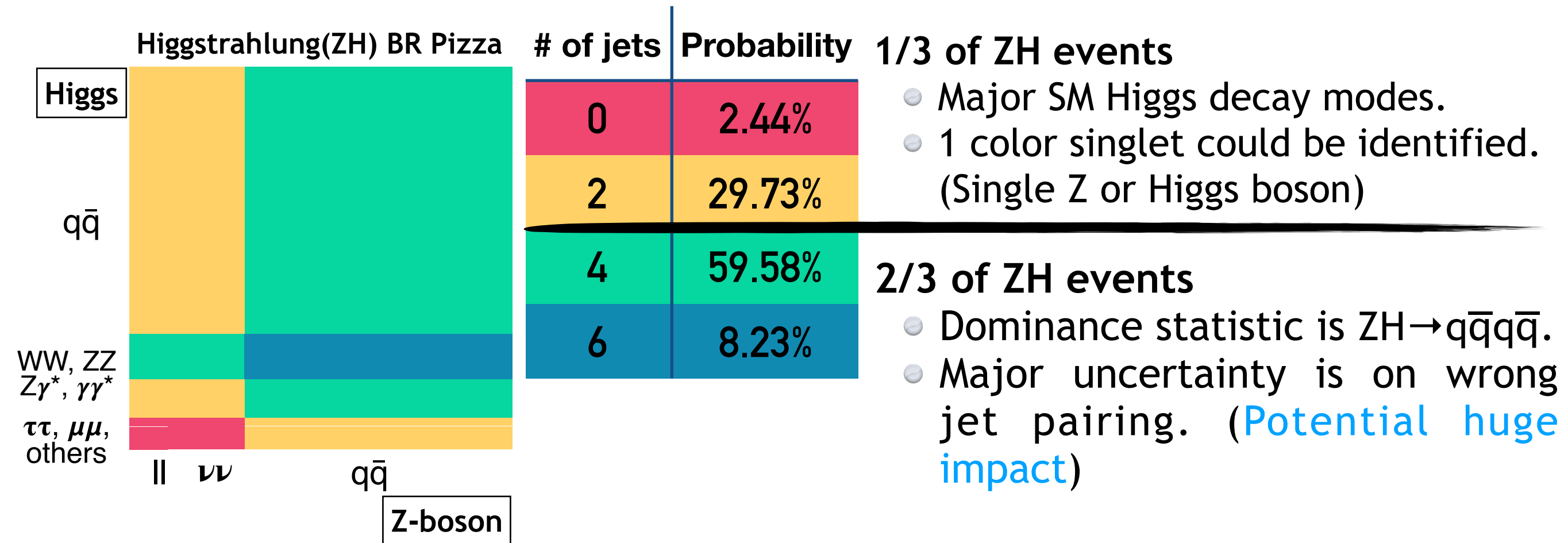
- Observables: Higgs mass, CP, $\sigma(ZH)$, event rate ($\sigma(ZH, \nu\nu H) \cdot \text{Br}(H \rightarrow X)$), Diff. distributions
 → Absolute Higgs width, branching ratio, couplings

Jets at the Higgs Signal



- Up to **97%** of Higgstrahlung(ZH) final-states associates to jets.
- Jets are also critical for many EW precision measurements.

Jets at the Higgs Signal



- 67% (4 + 6 jets) needs **dedicated color-singlet identification**: grouping the hadronic final-state particles into color-singlets (Z, W, H, γ^*). Can be done via jet clustering and pairing.
- **Jet clustering** is also essential for **differential** & EW precision measurements (e.g. TGCs).

# of jets	Probability	1/3 of ZH events	BMI: Massive bosons invariant mass resolutions
0	2.44%	<ul style="list-style-type: none"> Major SM Higgs decay modes. 1 color singlet could be identified. 	
2	29.73%	(Single Z or Higgs boson)	
4	59.58%	2/3 of ZH events	
6	8.23%	<ul style="list-style-type: none"> Dominance statistic of $ZH \rightarrow q\bar{q}q\bar{q}$. Major uncertainty is on wrong jet pairing. (Potential huge impact) 	

- 67% (4 + 6 jets) needs **dedicated color-singlet identification**: grouping the hadronic final-state particles into color-singlets (Z, W, H, γ^*). Can be done via jet clustering and pairing.
- **Jet clustering** is also essential for **differential** & EW precision measurements (e.g. TGCs).

BMI: Massive bosons

invariant mass resolutions

# of jets	Probability
0	2.44%
2	29.73%
4	59.58%
6	8.23%

1/3 of ZH events

- Major SM Higgs decay modes.
- 1 color singlet could be identified.
(Single Z or Higgs boson)

2/3 of ZH events

- Dominance statistic of $ZH \rightarrow q\bar{q}q\bar{q}$.
- Major uncertainty is on wrong jet pairing. (Potential huge impact)

BM2: # of jet identification & thrust clustering method for 2 jets

- 67% (4 + 6 jets) needs **dedicated color-singlet identification**: grouping the hadronic final-state particles into color-singlets (Z, W, H, γ^*). Can be done via jet clustering and pairing.
- Jet clustering** is also essential for **differential** & EW precision measurements (e.g. TGCs).

BMI: Massive bosons
invariant mass resolutions

# of jets	Probability	1/3 of ZH events
0	2.44%	<ul style="list-style-type: none"> Major SM Higgs decay modes. 1 color singlet could be identified.
2	29.73%	(Single Z or Higgs boson)
4	59.58%	2/3 of ZH events
6	8.23%	<ul style="list-style-type: none"> Dominance statistic of $ZH \rightarrow q\bar{q}q\bar{q}$. Major uncertainty is on wrong jet pairing. (Potential huge impact)

BM2: # of jet identification & thrust clustering method for 2 jets

- 67% (4 + 6 jets) needs **dedicated color-singlet identification**: grouping the hadronic final-state particles into color-singlets (Z, W, H, γ^*). Can be done via jet clustering and pairing.
- **Jet clustering** is also essential for **differential & EW precision** measurements (e.g. TGCs).

BM3: Jet energy and angular differential response

# of jets	Probability
0	2.44%
2	29.73%
4	59.58%
6	8.23%

1/3 of ZH events

- Major SM Higgs decay modes.
- 1 color singlet could be identified.
(Single Z or Higgs boson)

BMI: Massive bosons

invariant mass resolutions

2/3 of ZH events

- Dominance statistic of $ZH \rightarrow q\bar{q}q\bar{q}$.
- Major uncertainty is on wrong jet pairing. (Potential huge impact)

BM2: # of jet identification & thrust clustering method for 2 jets

BM4: Separation of WW , ZZ , and ZH decay to $qqqq$ final state

- 67% (4 + 6 jets) needs **dedicated color-singlet identification**: grouping the hadronic final-state particles into color-singlets (Z, W, H, γ^*). Can be done via jet clustering and pairing.
- Jet clustering** is also essential for **differential & EW precision** measurements (e.g. TGCs).

BM3: Jet energy and angular differential response

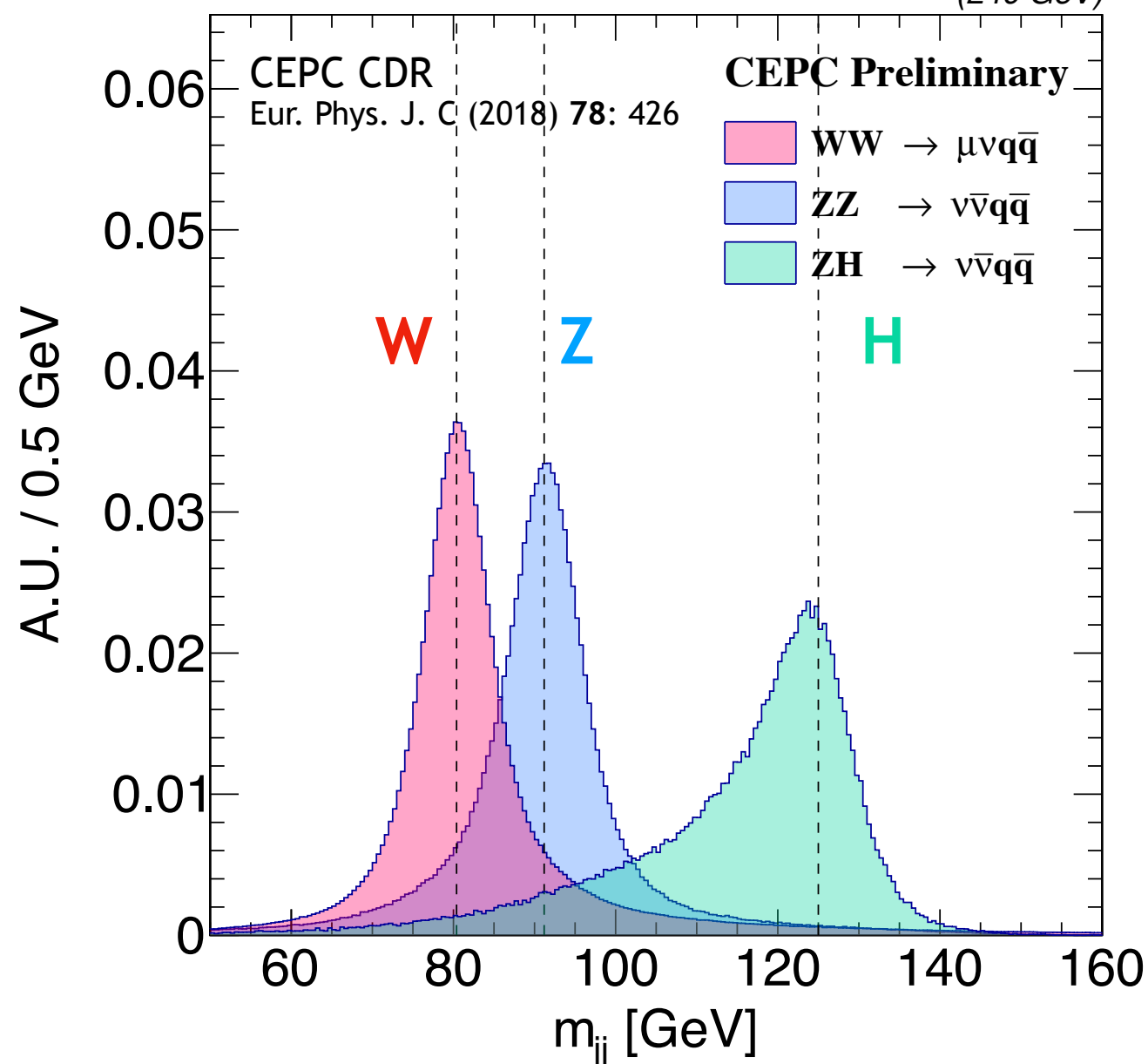
# of jets	Probability	
0	2.44%	1/3 of ZH events BMI: Massive bosons invariant mass resolutions <ul style="list-style-type: none"> Major SM Higgs decay modes. 1 color singlet could be identified. (Single Z or Higgs boson)
2	29.73%	
4	59.58%	2/3 of ZH events <ul style="list-style-type: none"> Dominance statistic of $ZH \rightarrow q\bar{q}q\bar{q}$. Major uncertainty is on wrong jet pairing. (Potential huge impact)
6	8.23%	

- 67% (4 + 6 jets) needs dedicated color-singlet identification: grouping the hadronic final-state particles into color-singlets (Z, W, H, γ^*). Can be done via jet clustering and pairing.
- Jet clustering is also essential for differential & EW precision measurements (e.g. TGCs).

BM1: Massive Boson Mass Resolution

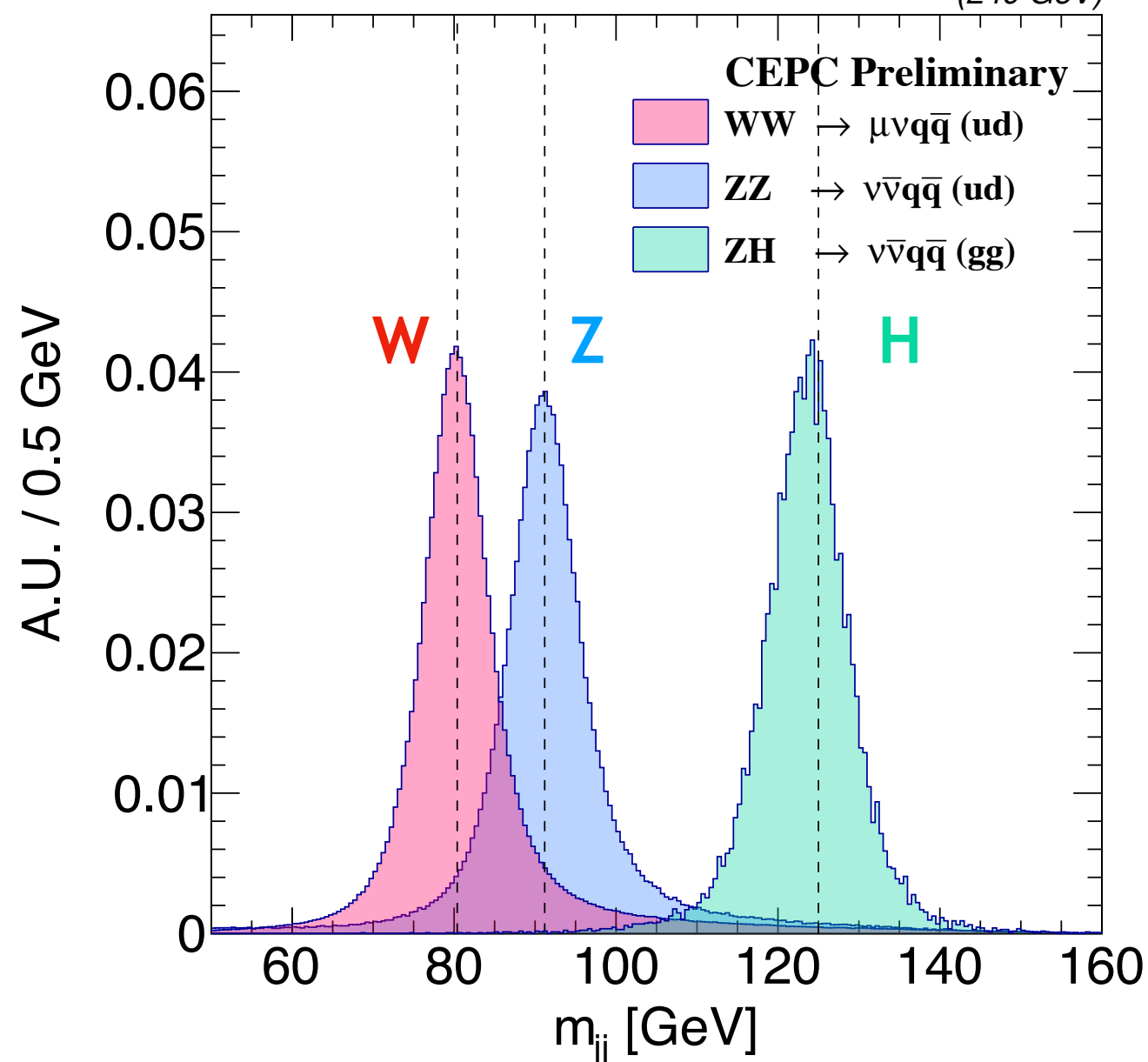
Before Cleaned

(240 GeV)



After Cleaned

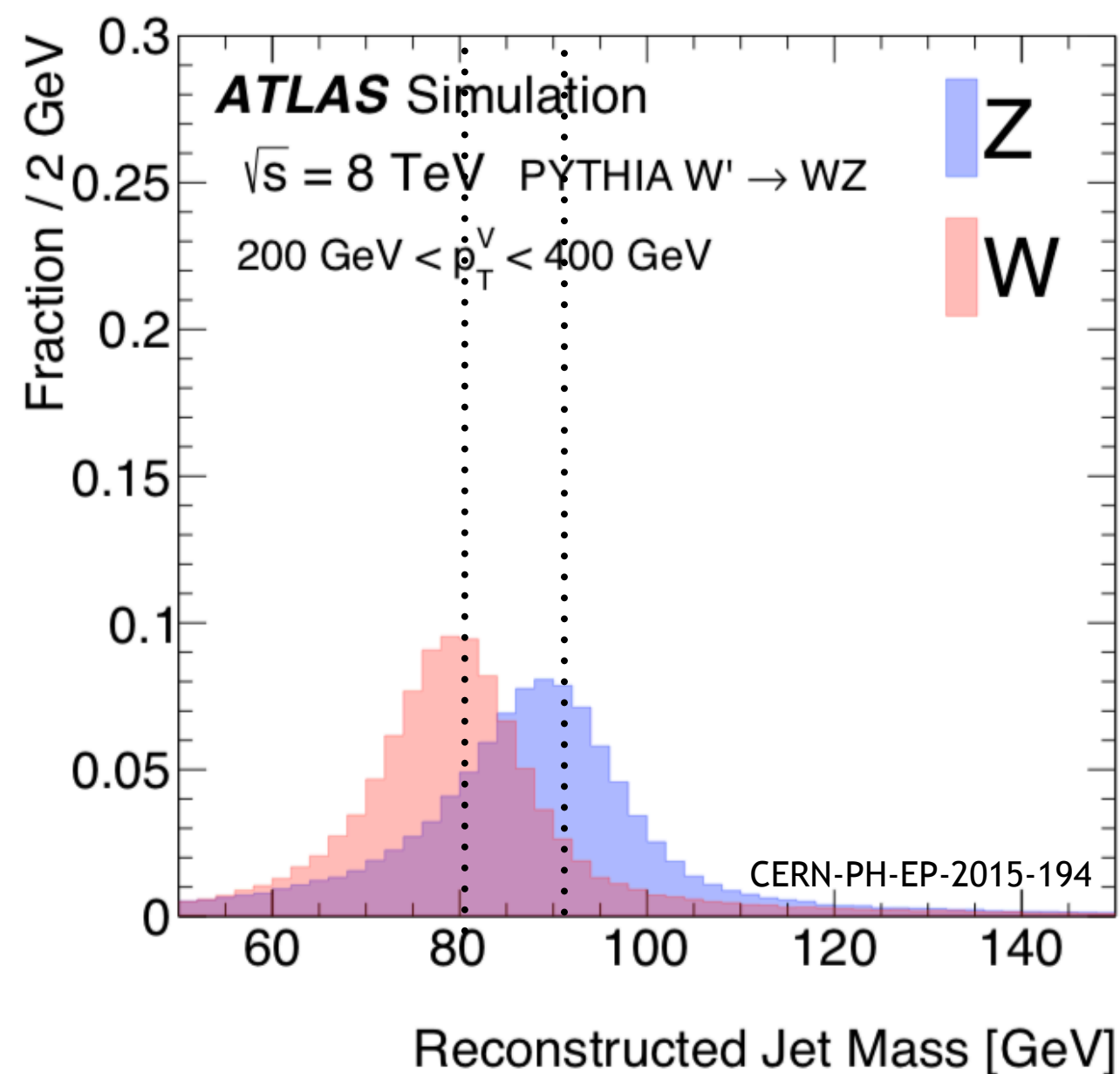
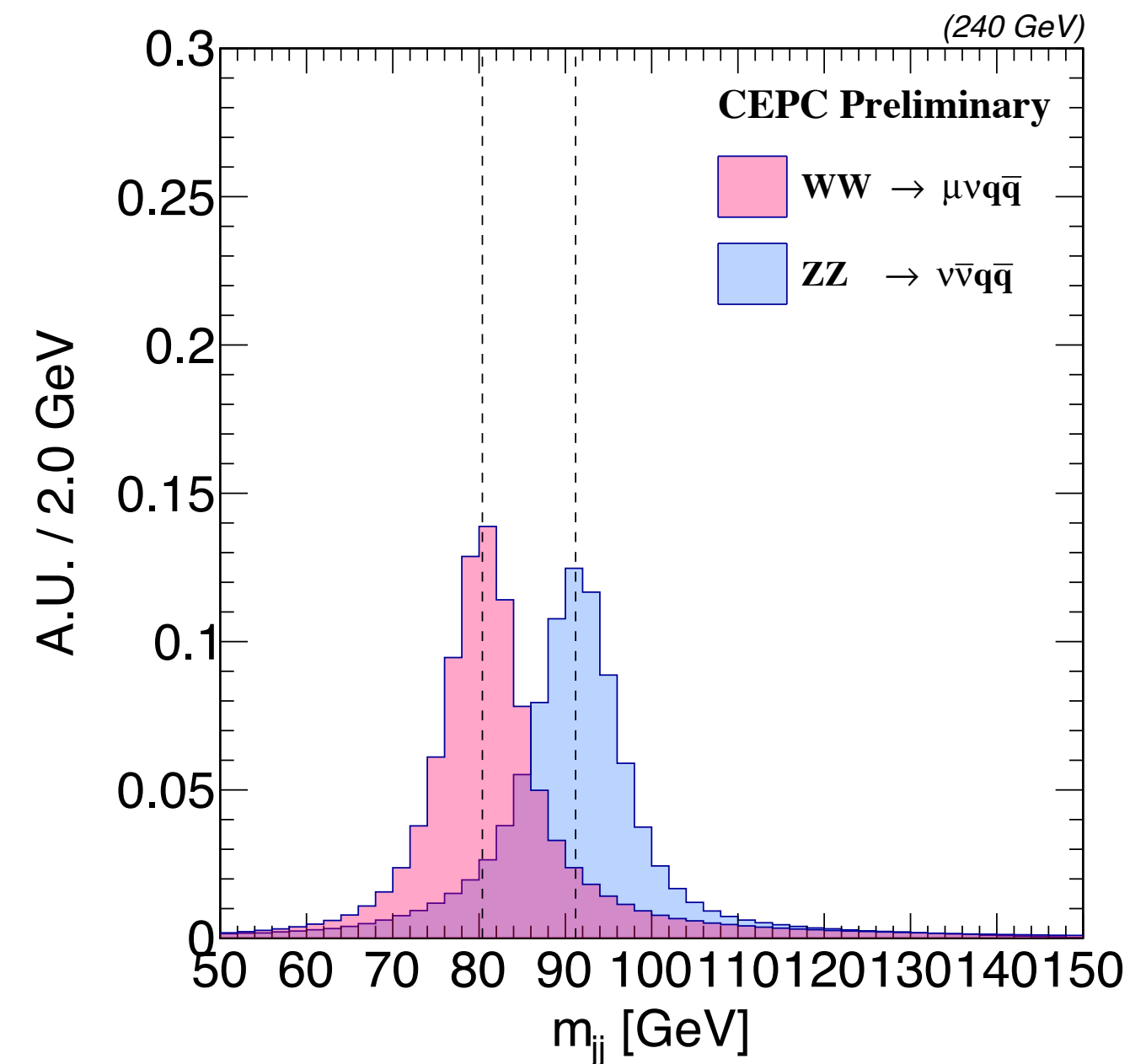
(240 GeV)



- **W-, Z-, and Higgs-boson masses in dijet final state can be well separated at CEPC.**
- After cleaned, Z- and W-boson could be separated $\approx 2\sigma$, and the Higgs **Boson Mass Resolution = 3.8%** achieving the CEPC baseline.

Cleaned: Select the light flavor jet event with low energy ISR, low energy neutrino inside jet, and within $|\cos\theta| < 0.85$.

BM1: Massive Boson Mass Resolution



- The separation of Z- and W-boson at CEPC is much better than ATLAS as it should be, because of the better collision environment and detector response.

# of jets	Probability	1/3 of ZH events
0	2.44%	<ul style="list-style-type: none"> Major SM Higgs decay modes. 1 color singlet could be identified. (Single Z or Higgs boson)
2	29.73%	
4	59.58%	2/3 of ZH events
6	8.23%	<ul style="list-style-type: none"> Dominance statistic of $ZH \rightarrow q\bar{q}q\bar{q}$. Major uncertainty is on wrong jet pairing. (Potential huge impact)

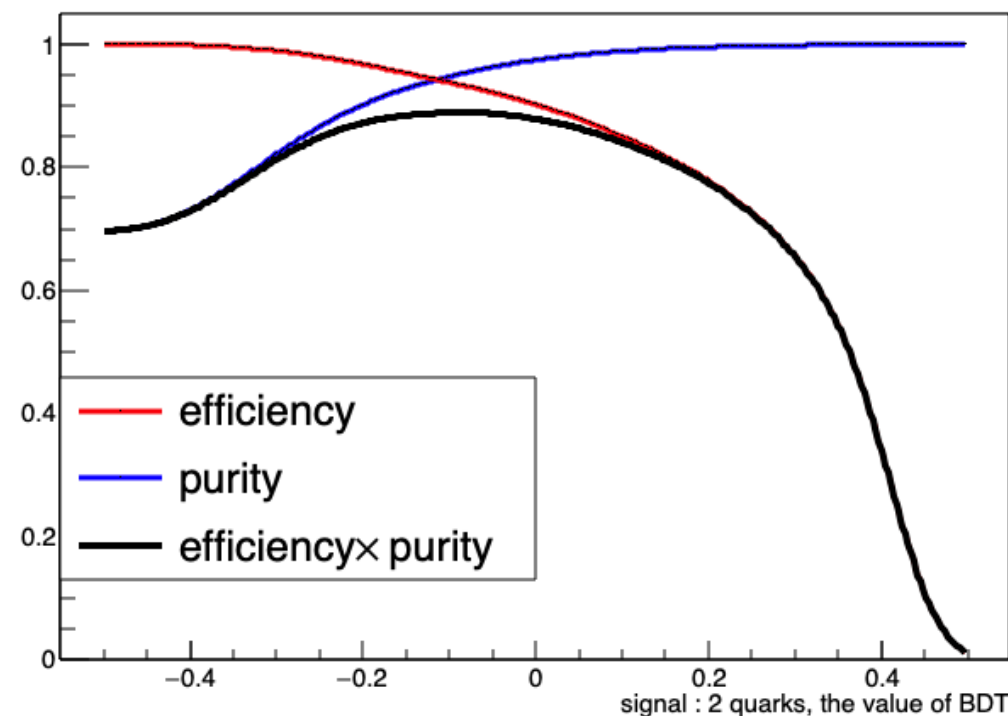
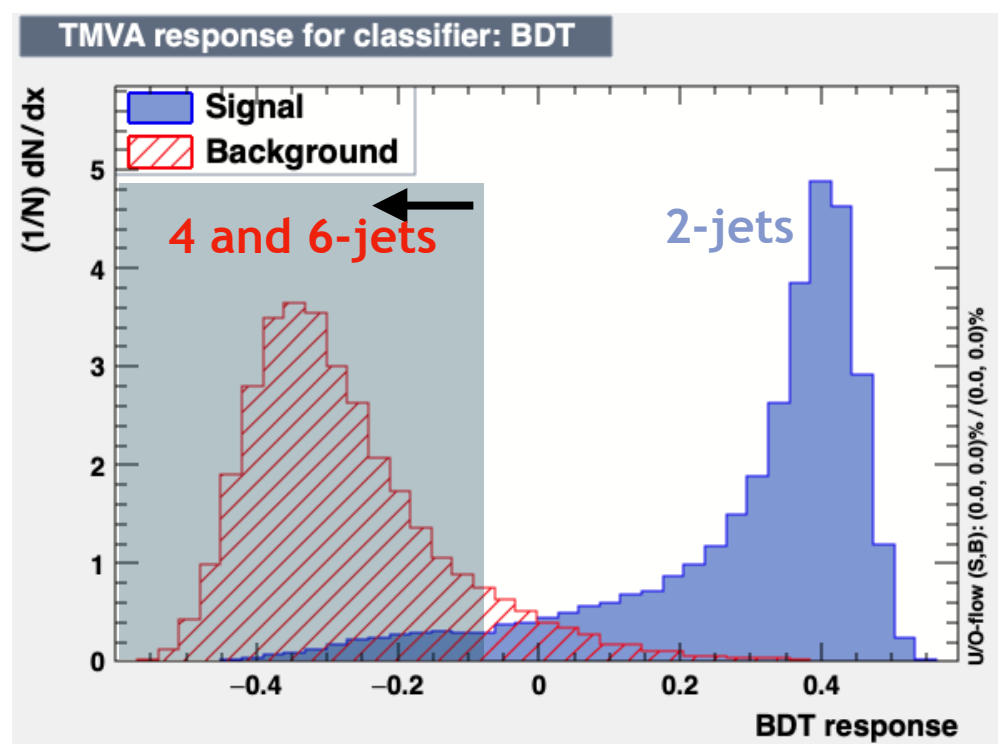
BM2: # of jet identification & thrust clustering method for 2 jets

- 67% (4 + 6 jets) needs **dedicated color-singlet identification**: grouping the hadronic final-state particles into color-singlets (Z, W, H, γ^*). Can be done via jet clustering and pairing.
- **Jet clustering** is also essential for **differential** & EW precision measurements (e.g. TGCs).



BM2: Preliminary Number of Jet Identification

Yong-Feng Zhu



Samples:

$e^+e^- \rightarrow q\bar{q}$ (2 jets)
 $ZZ \rightarrow q\bar{q}q\bar{q}$ (4 jets)
 $W^+W^- \rightarrow q\bar{q}q\bar{q}$ (4 jets)
 $ZH \rightarrow q\bar{q}q\bar{q}$ (4 jets)
 $ZH \rightarrow q\bar{q}H \rightarrow qq\bar{q}\bar{q}qq$ (6 jets)

Signal	<i>Efficiency × purity</i>
2 jets	88.4%
6 jets	1.8%

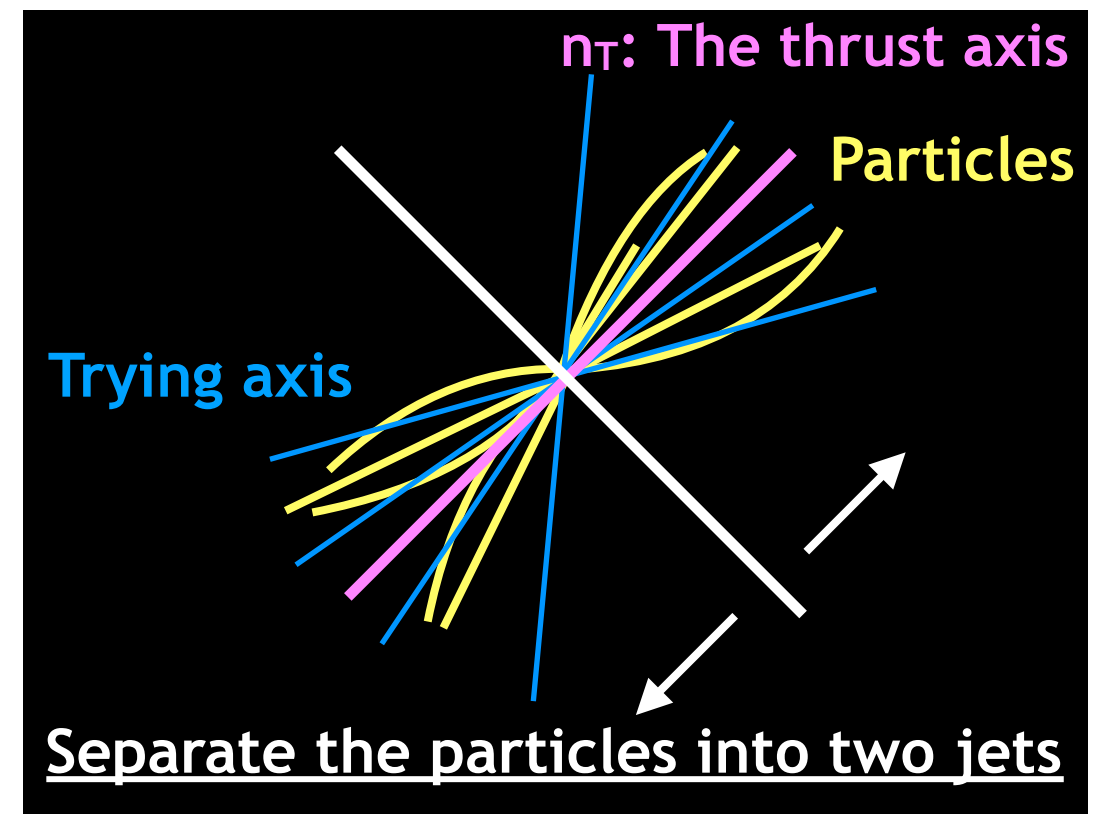
20 event-shape variables are combined with the multi-variate analysis to separate 2, 4, and 6 jets final-states.

BM2: Thrust Jet Clustering Method

$$T \equiv \max \frac{\sum_j^N |\underline{P_j} \cdot \underline{n_T}|}{\sum_i^N |\underline{P_i}|}$$

$\underline{P_i}$ or $\underline{P_j}$: Momentum of each particle

$\underline{n_T}$: A unit vector ($\sin \theta \times \cos \phi$, $\sin \theta \times \sin \phi$, $\cos \theta$)



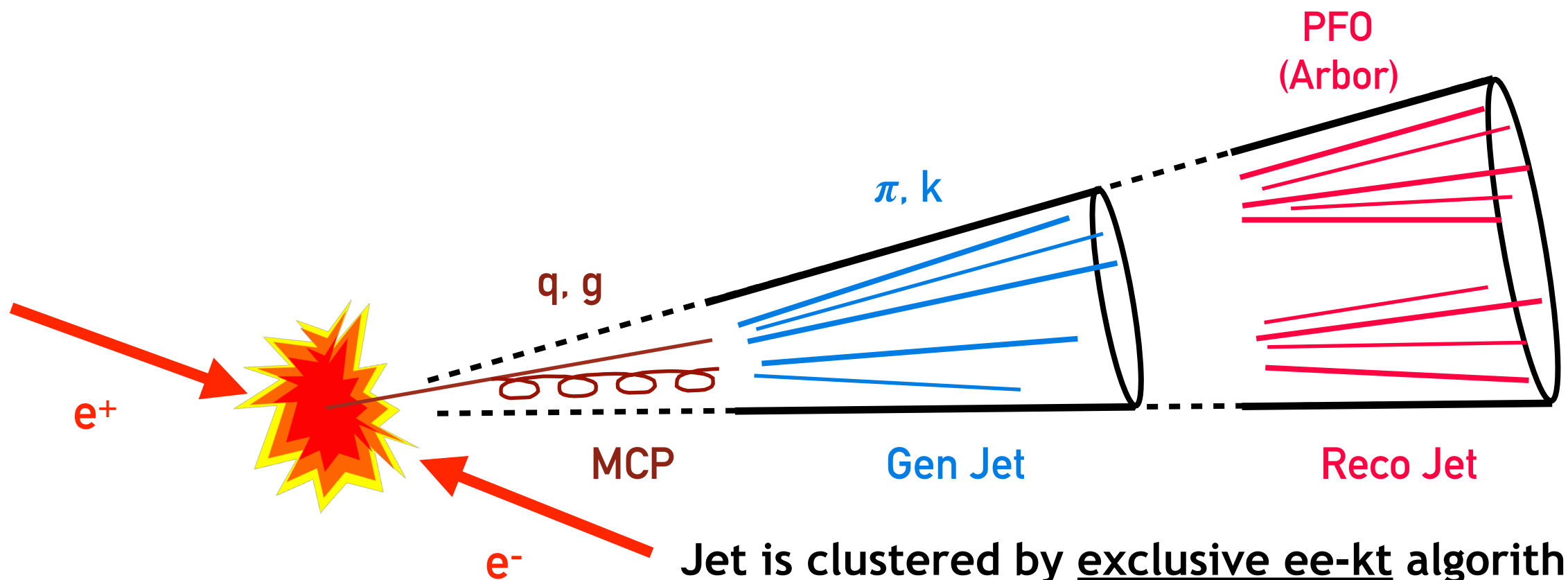
- “Thrust” is one kind of event-shape variables.
- The nature clustering idea for the **single boson decays to di-jet events**, thrust.
 1. First, boost the system back to the rest frame.
 2. Find out a vector in the θ and ϕ phase space which has highest momentum flux.
 3. System is divided into 2 hemispheres with the thrust axis, and each identified as a jets. (Only applicable to **2 jets final-state**)

# of jets	Probability	1/3 of ZH events
0	2.44%	<ul style="list-style-type: none"> Major SM Higgs decay modes. 1 color singlet could be identified. (Single Z or Higgs boson)
2	29.73%	
4	59.58%	2/3 of ZH events <ul style="list-style-type: none"> Dominance statistic of $ZH \rightarrow q\bar{q}q\bar{q}$. Major uncertainty is on wrong jet pairing. (Potential huge impact)
6	8.23%	

- 67% (4 + 6 jets) needs **dedicated color-singlet identification**: grouping the hadronic final-state particles into color-singlets (Z, W, H, γ^*). Can be done via jet clustering and pairing.
- **Jet clustering** is also essential for **differential & EW precision measurements** (e.g. TGCs).

BM3: Jet energy and angular differential response

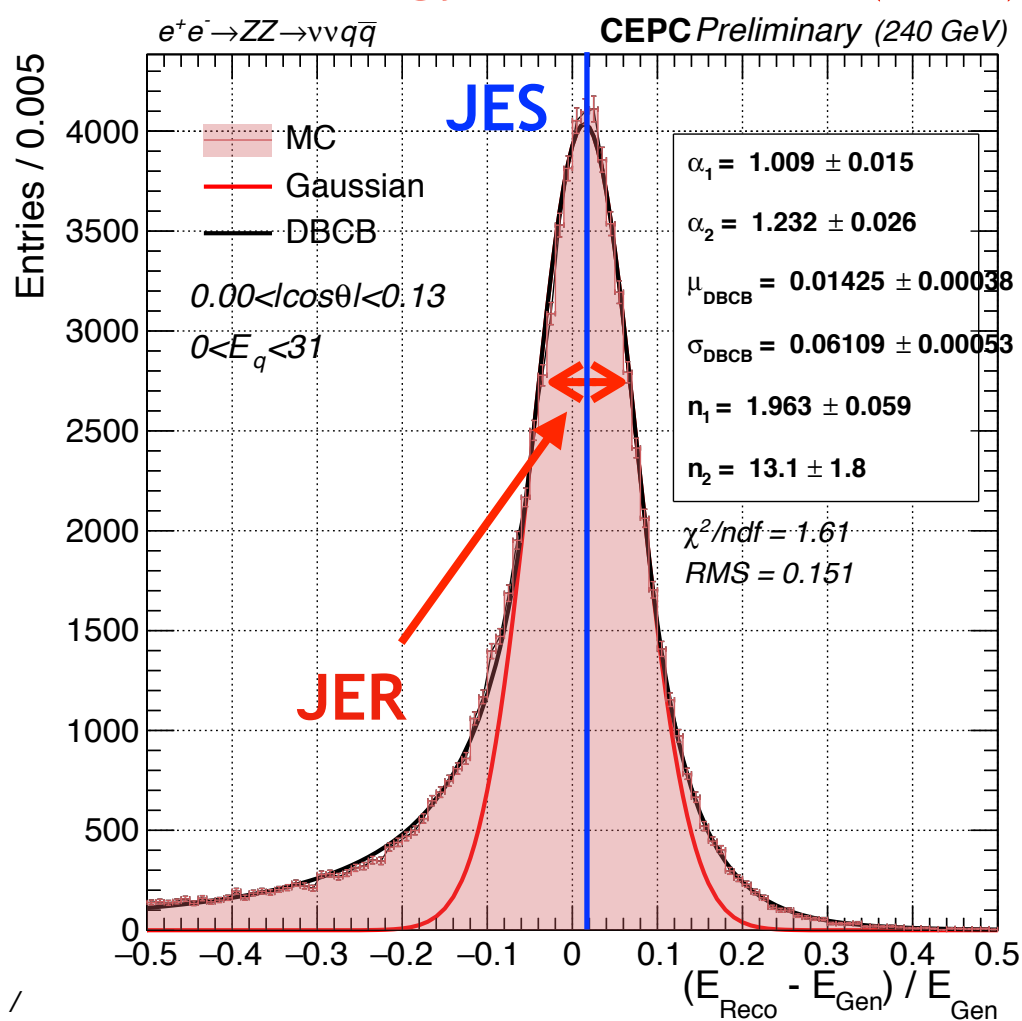
- **MCP** represents initial parton of MC quark. The original state of quark.
- **GenJets** are grouped all MC particles except neutrinos with $c\tau > 1$ cm through exclusive ee-kt jet clustering algorithm.
- **RecoJets** are grouped with the particle flow objects by exclusive ee-kt jet clustering algorithm.



- Double-sided crystal ball(DBCB) function is used to extract energy and angular resolution and scale.

Jet Energy Scale (JES)

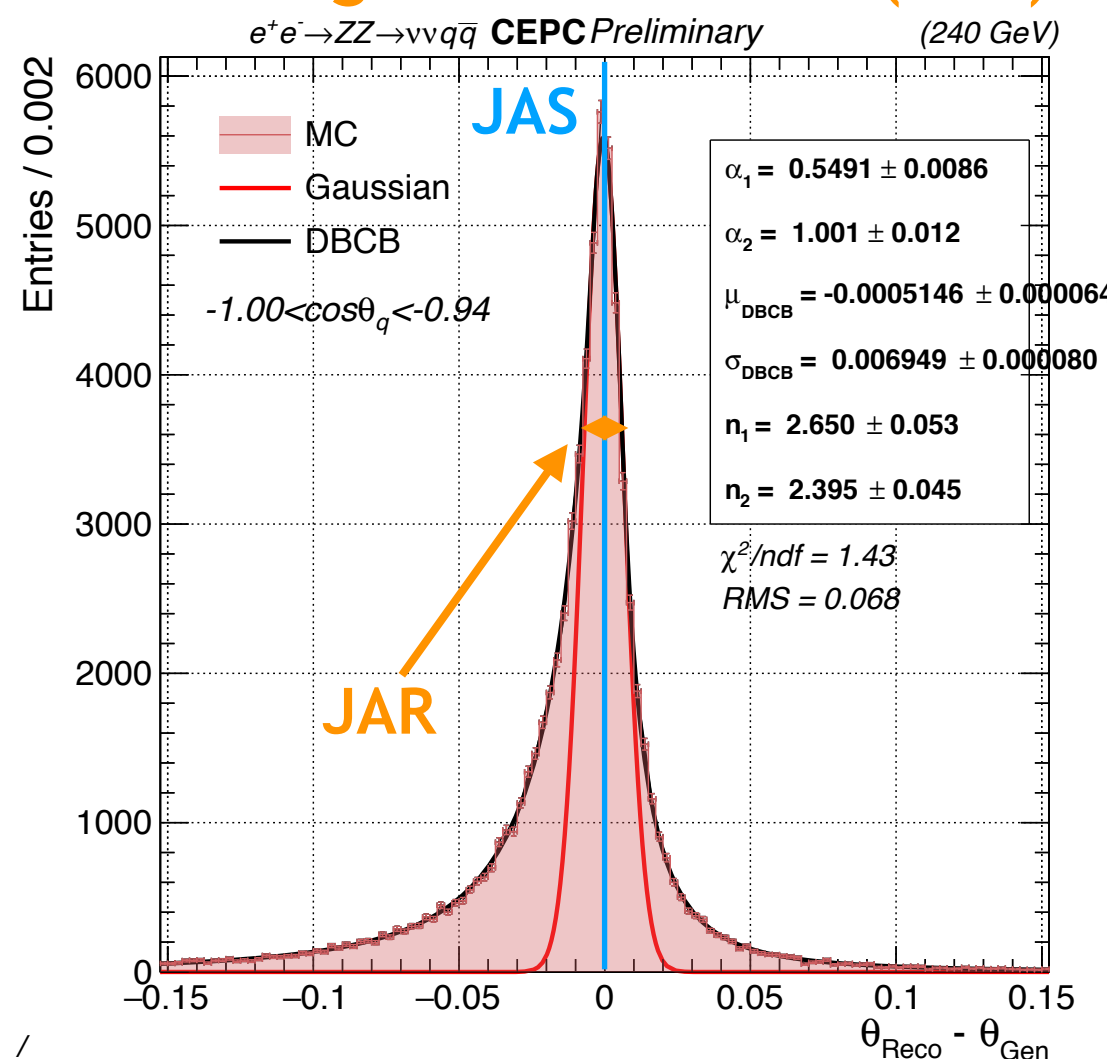
Jet Energy Resolution (JER)



Relative difference :
$$\frac{E_{\text{Reco}} - E_{\text{Gen}}}{E_{\text{Gen}}}$$

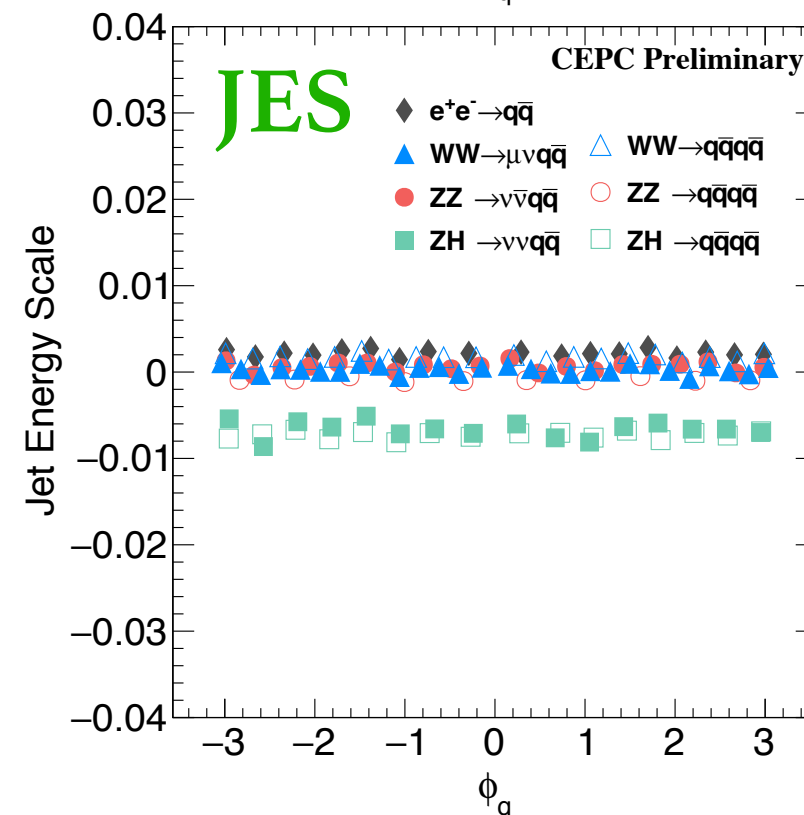
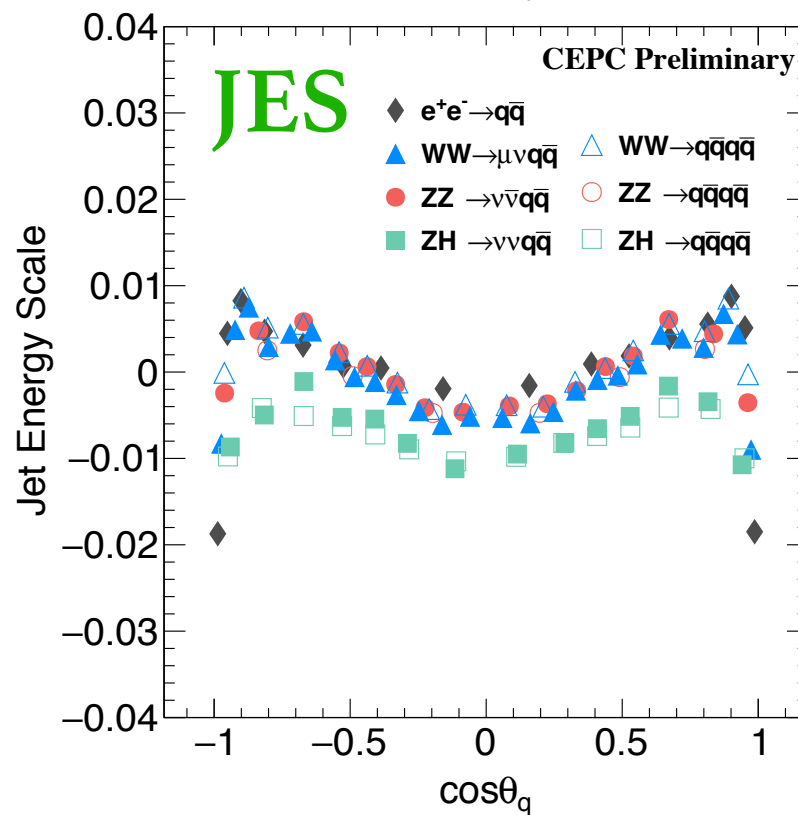
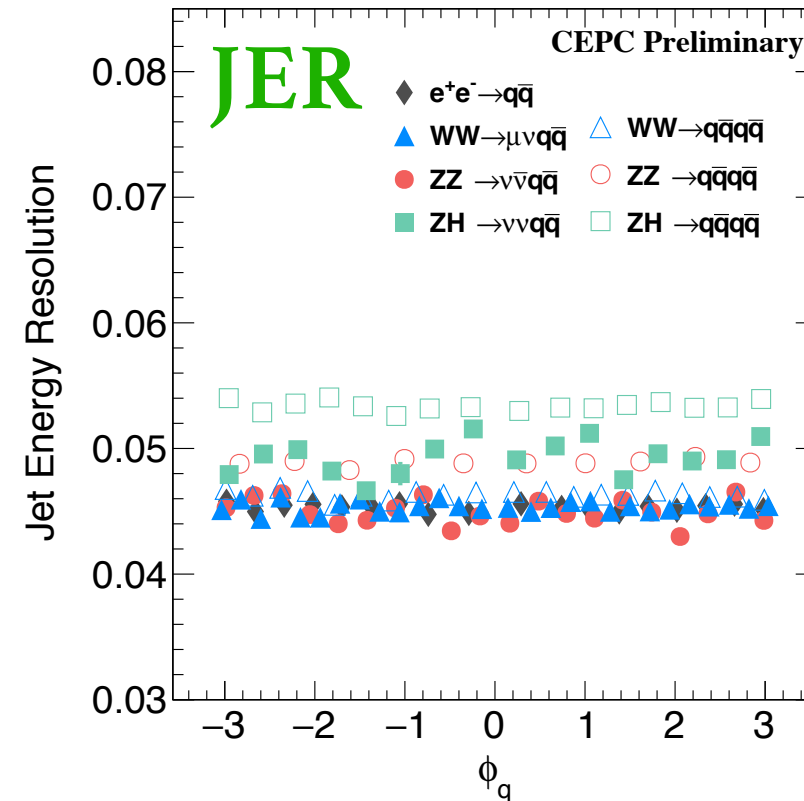
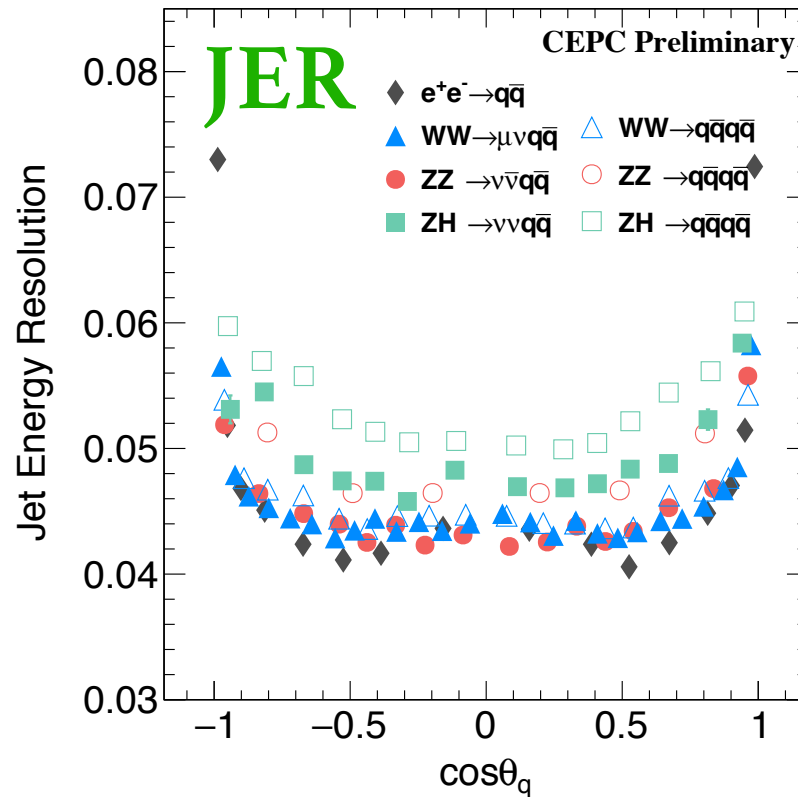
Jet Angular Scale (JAS)

Jet Angular Resolution (JAR)



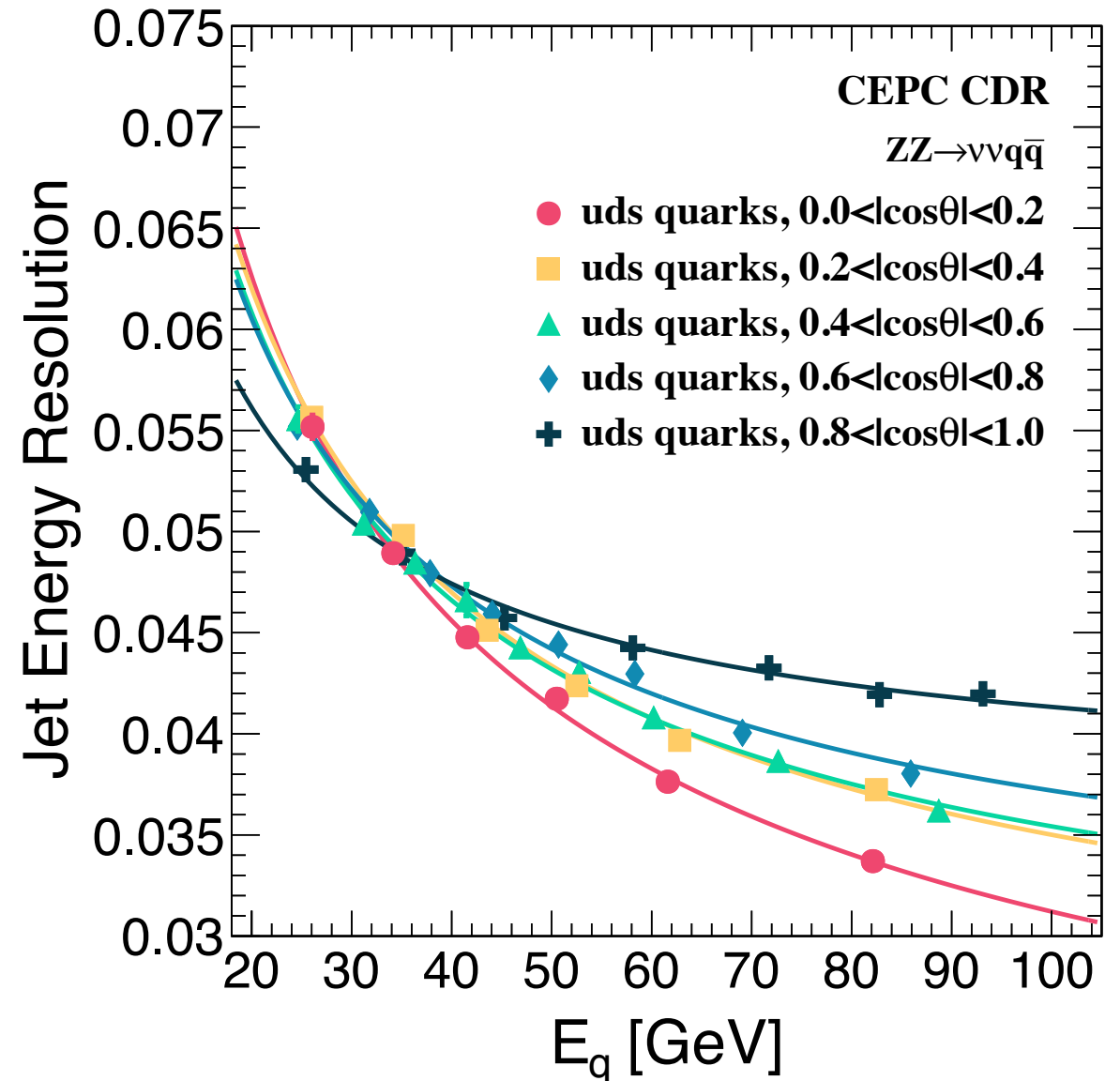
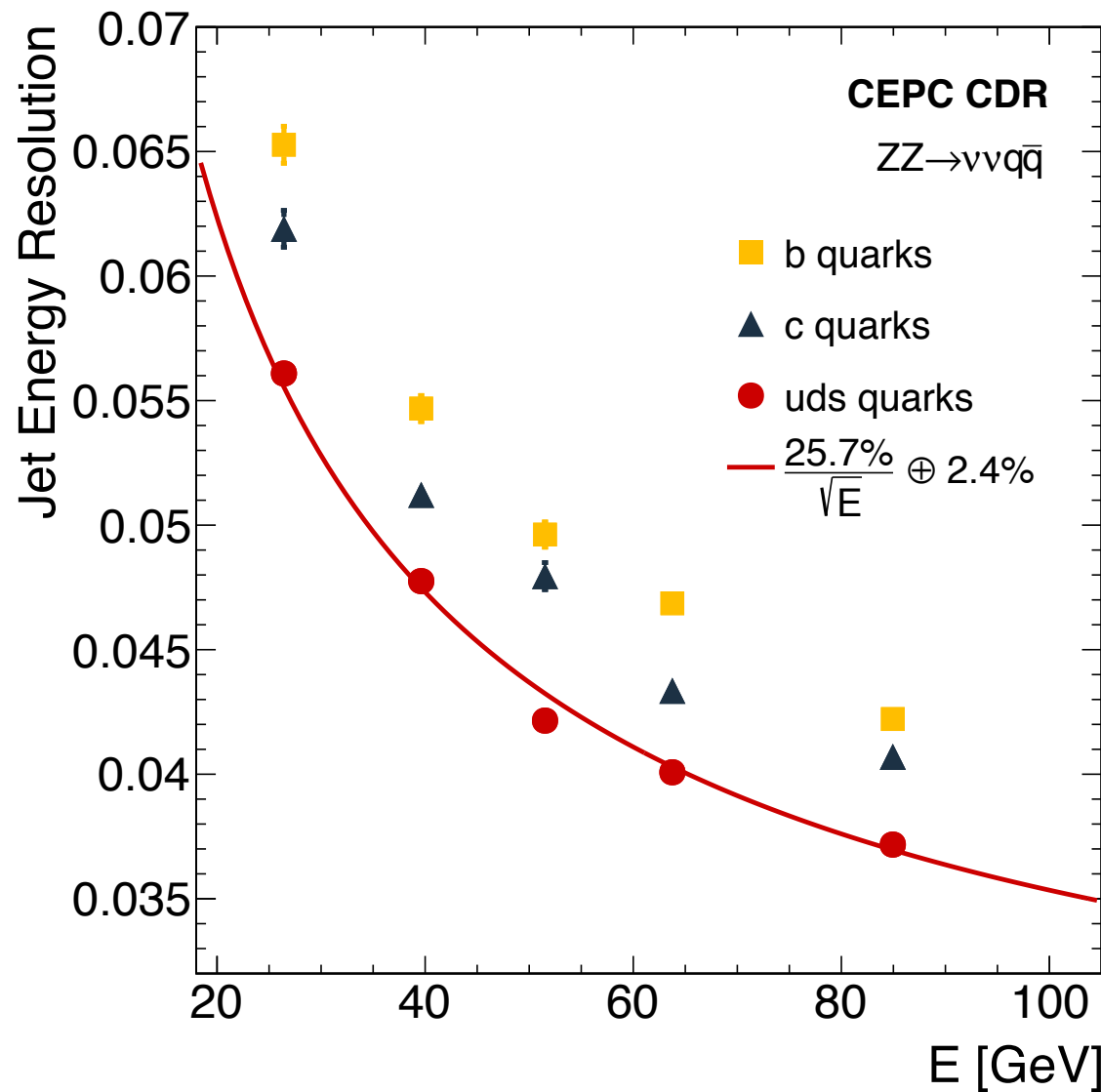
Difference :
$$\theta_{\text{Reco}} - \theta_{\text{Gen}}$$

BM3: JER & JES (Reco-Gen)



- JER is around **4.5%** in barrel region; JES is around **0**.
- The difference between 2 and 4 jets final-state is controlled within **1%** level.

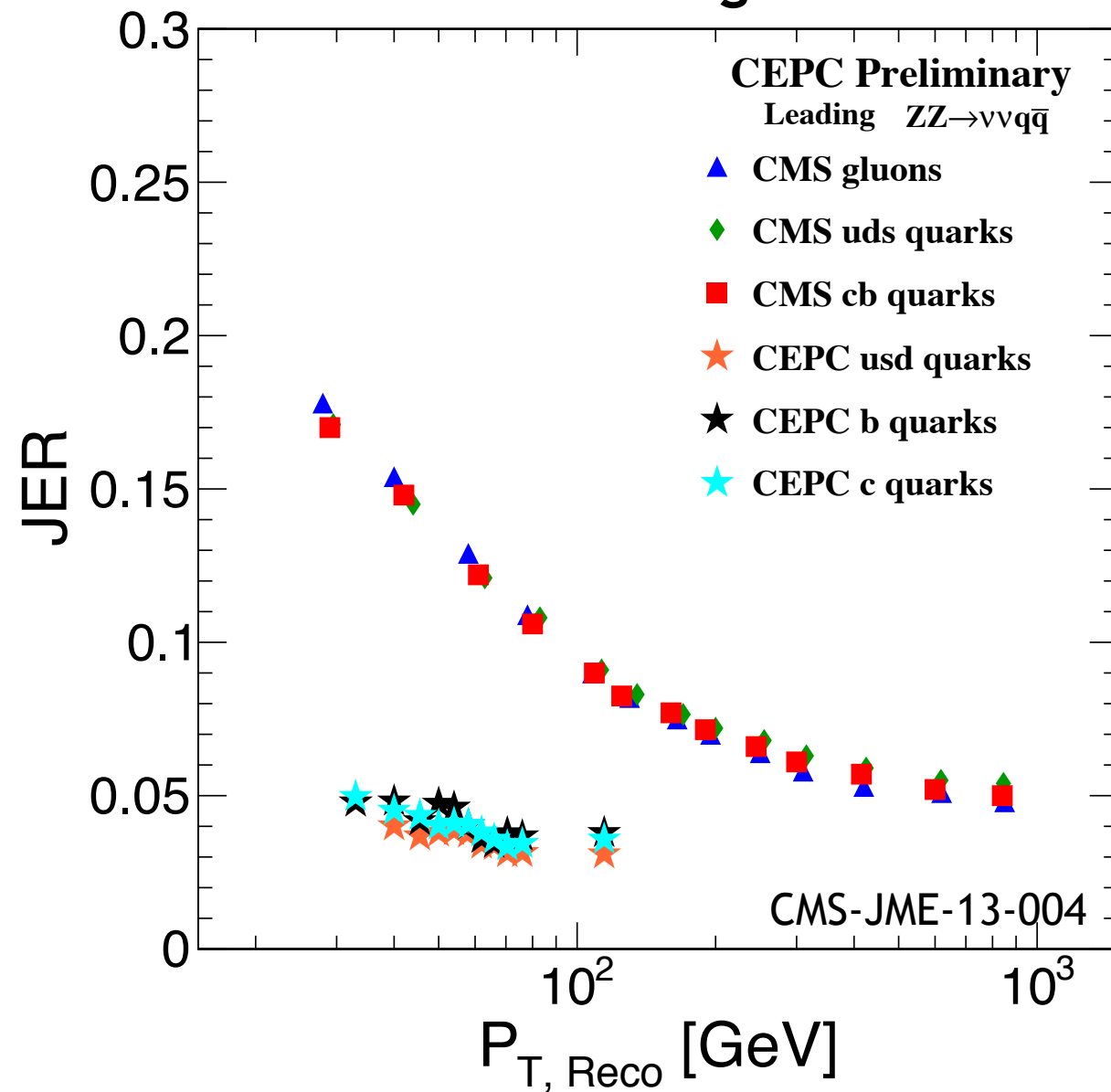
BM3: JER (Reco-Gen)



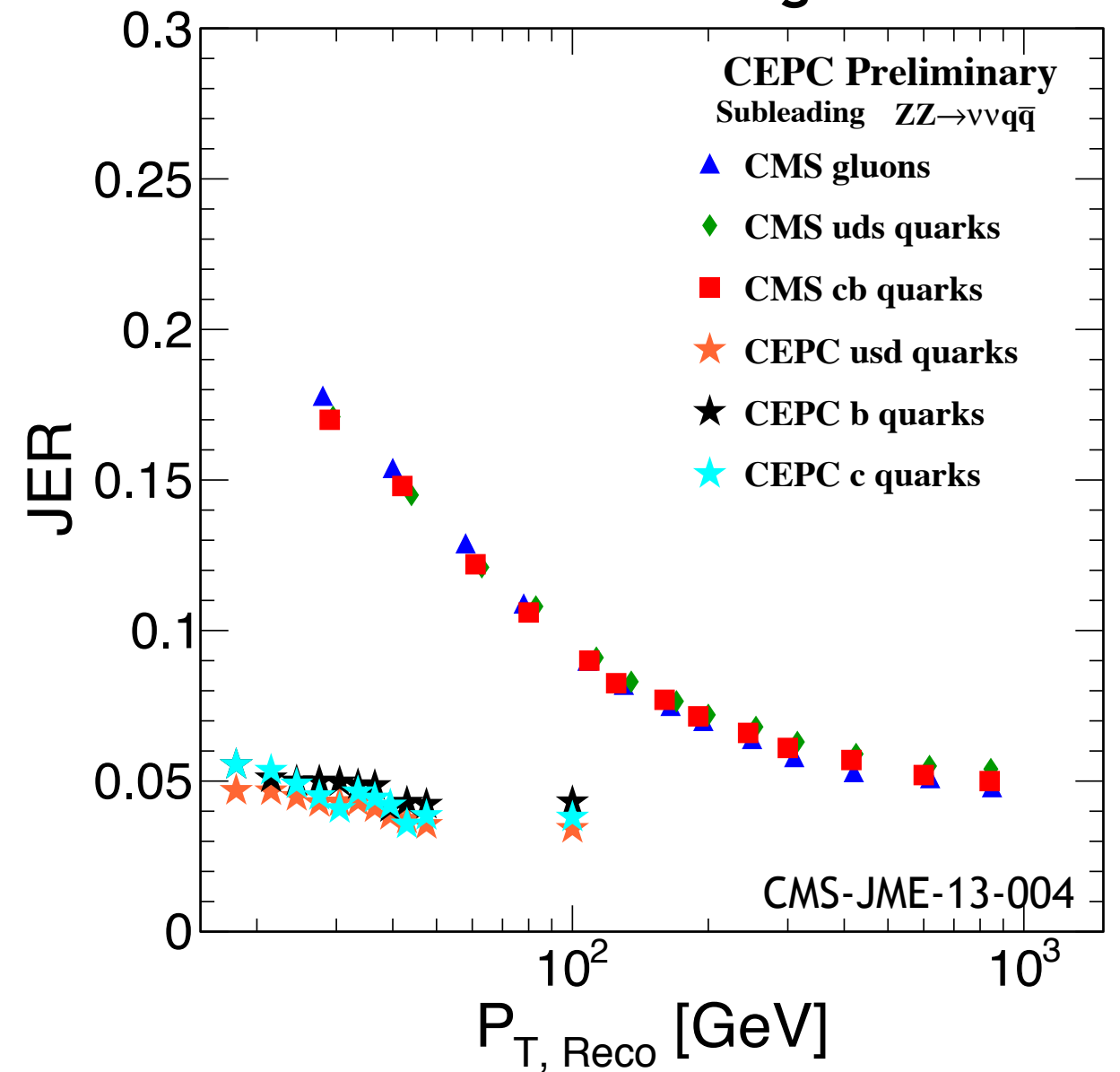
■ JER also depends on jet flavors.

■ For light-flavor jets with high energy and within central region of barrel, JER could reach **3%**.

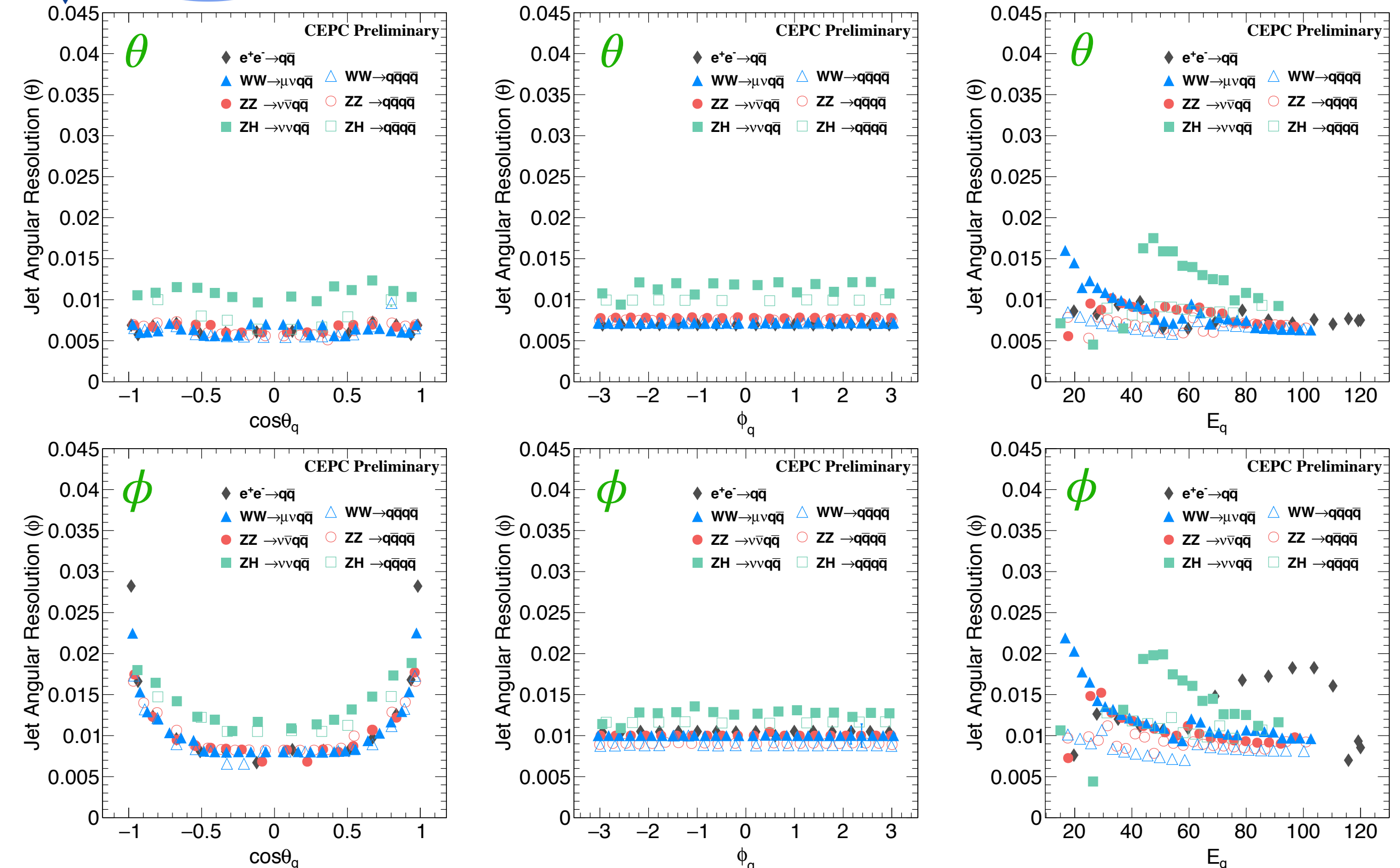
Leading



Sub-leading



- JER at CEPC is better than CMS as it should be; **2-4 times** better in the same energy region.



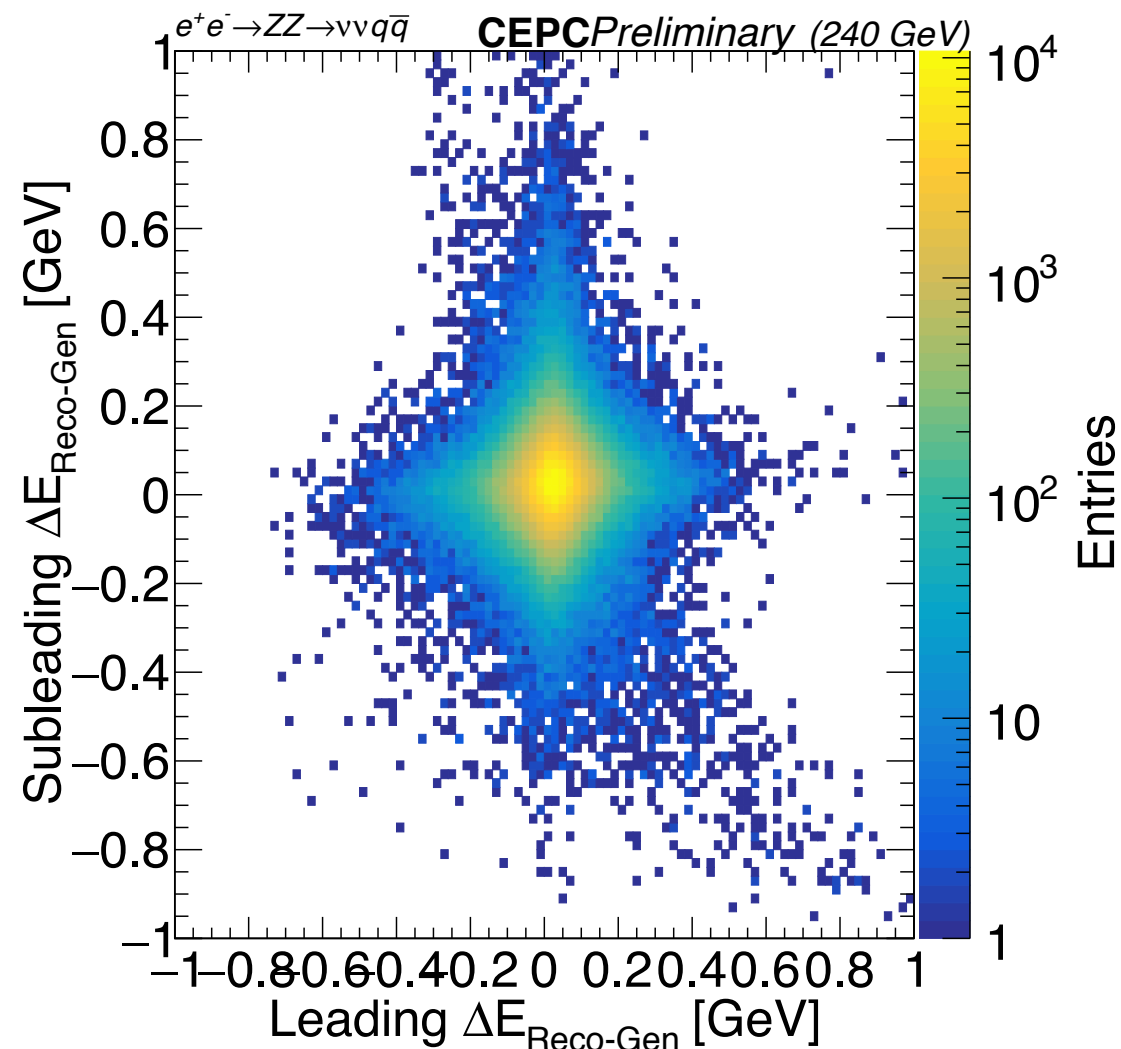
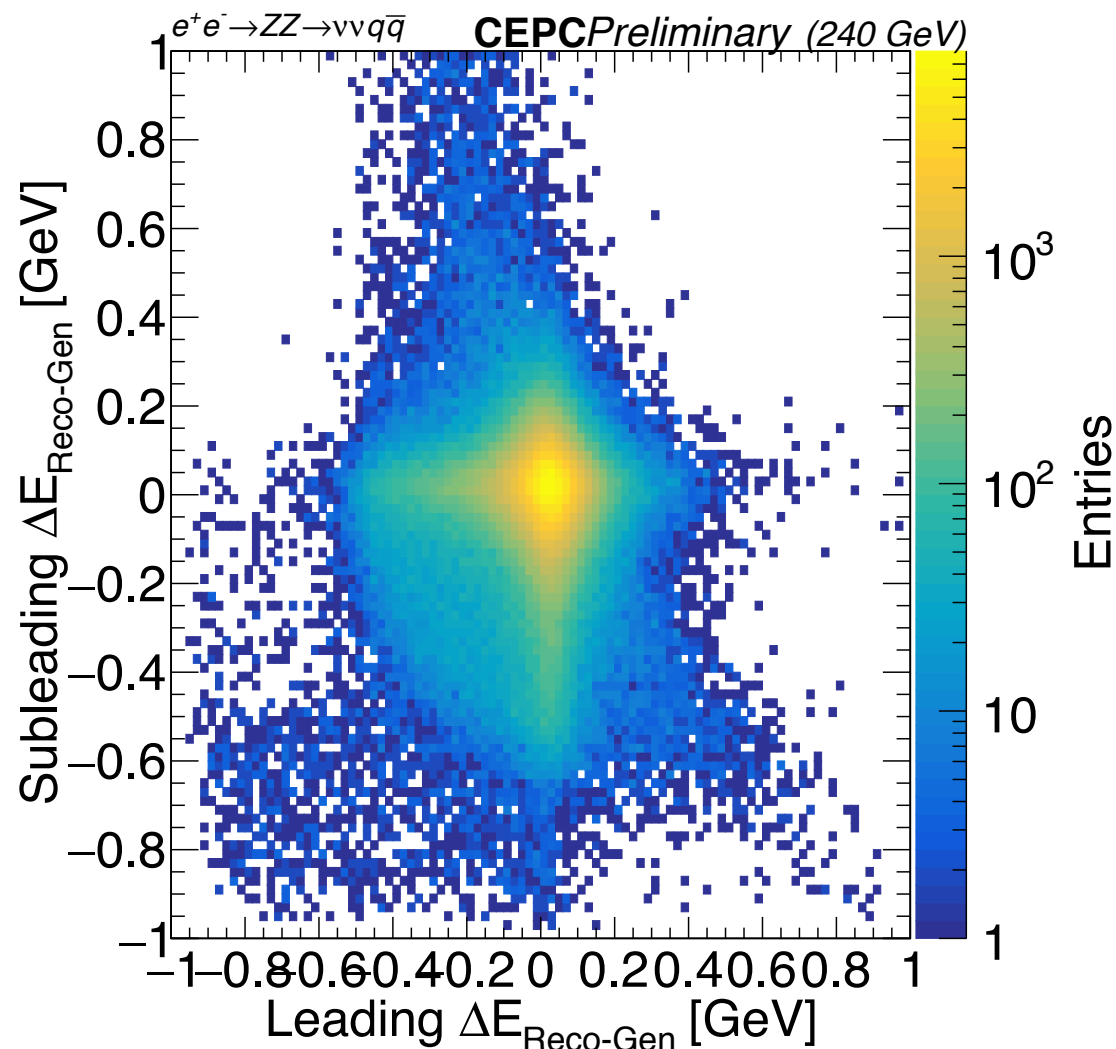
■ JAR is around **1%** in barrel region; JAS is **independent** of ϕ and energy.

■ The difference between 2 and 4 jets final-state is controlled within **1%** level.

BM3: Thrust Jet Clustering Method

ee-kt

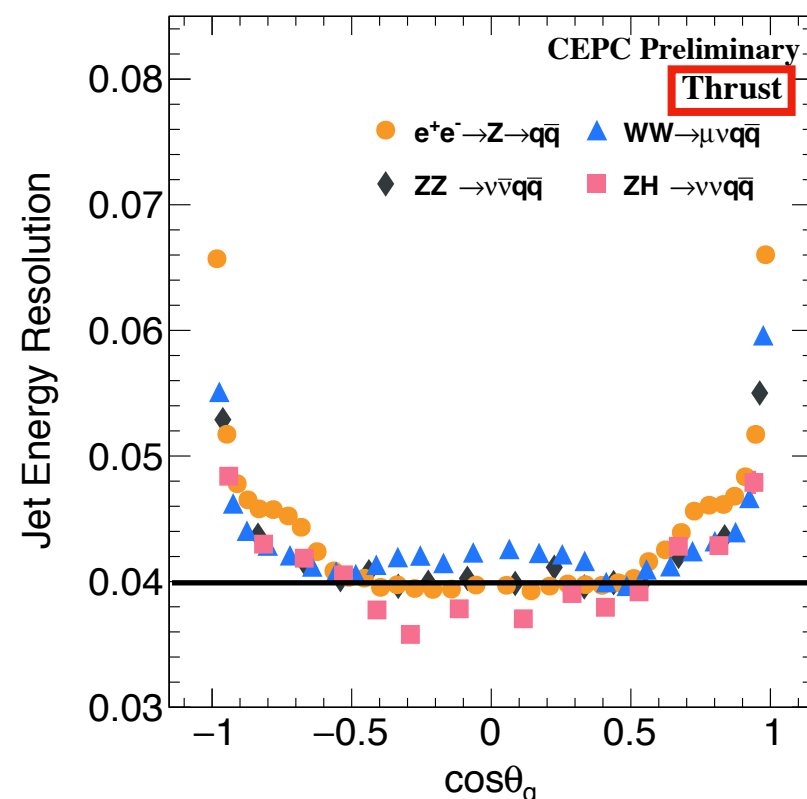
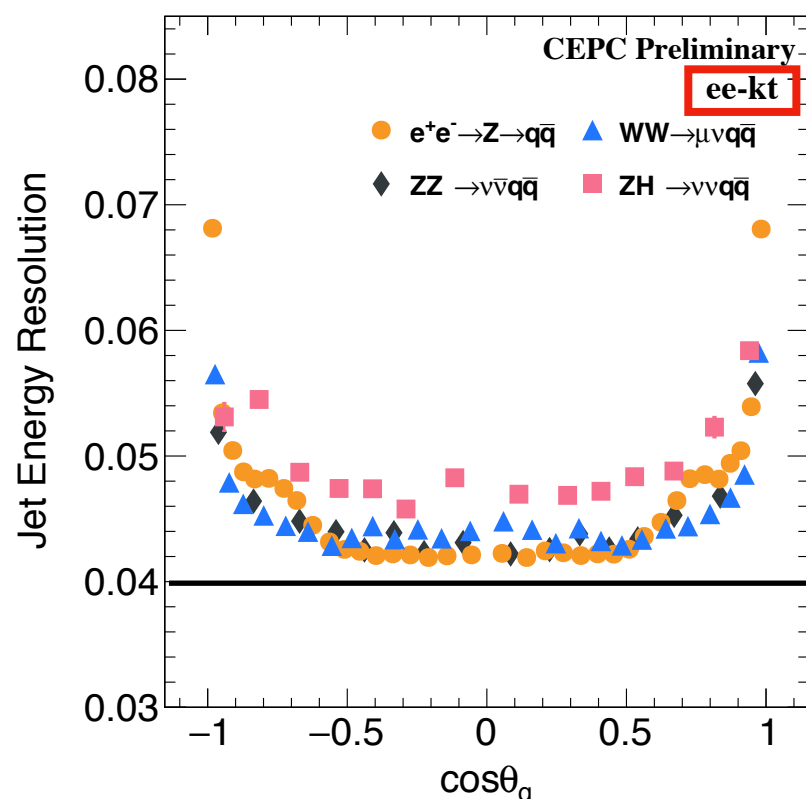
Thrust



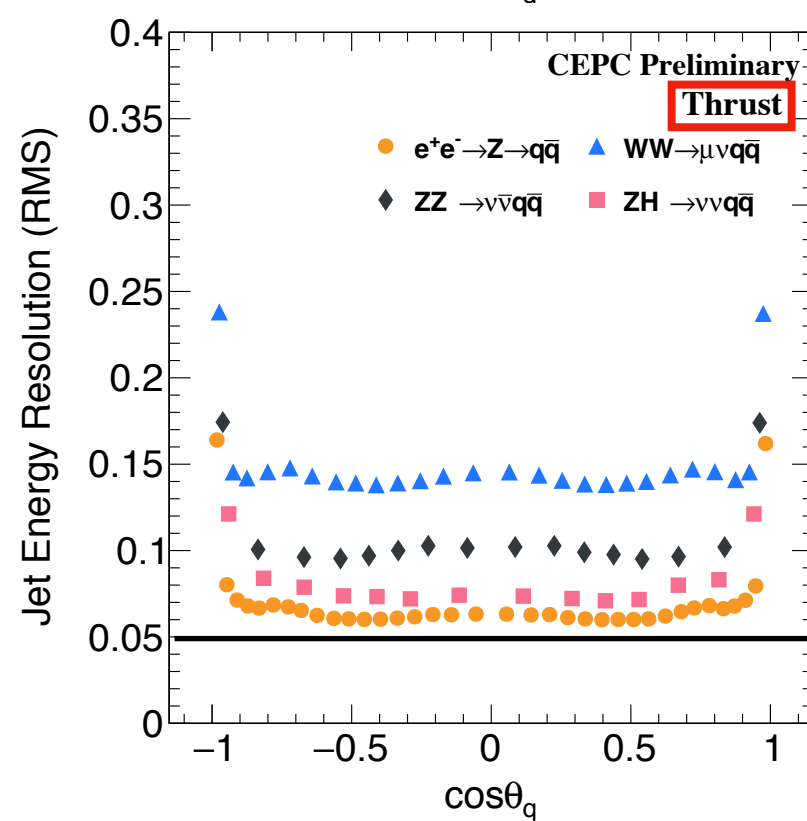
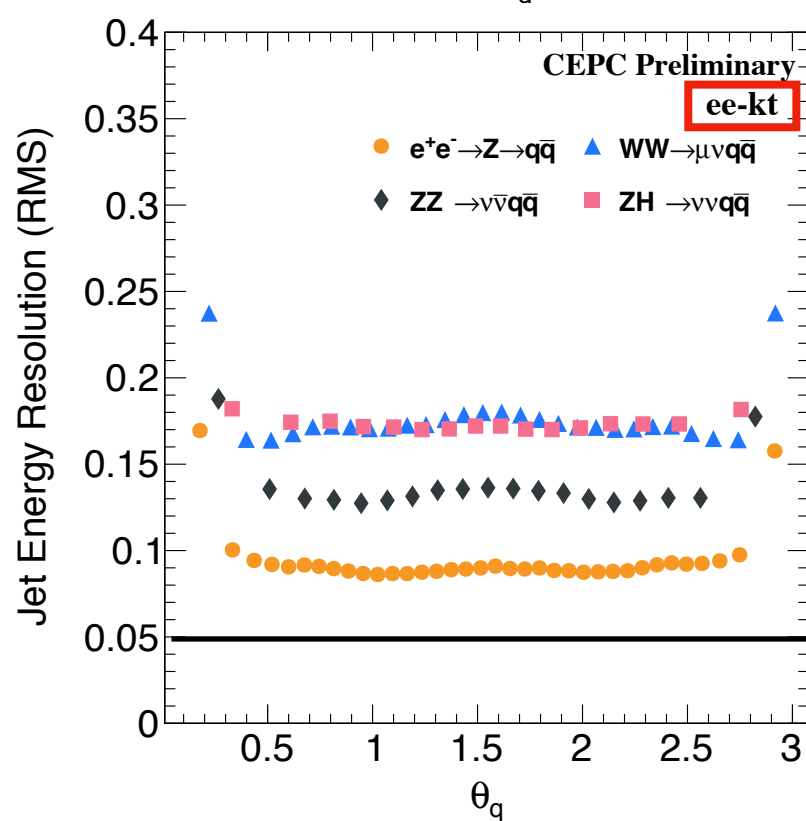
- Identify the 2 jets final-state event with (efficiency x purity) = **88.4%**, the thrust jet clustering method could be employed.
- After “cleaned” selection, the thrust method has **significant tail suppressed**
→ expected to have improvement on jet energy and angular response.

Cleaned: Select the light flavor jet event with low energy ISR, low energy neutrino inside jet, and within $|\cos\theta| < 0.85$.

BM3: JER (ee-kt—Thrust)



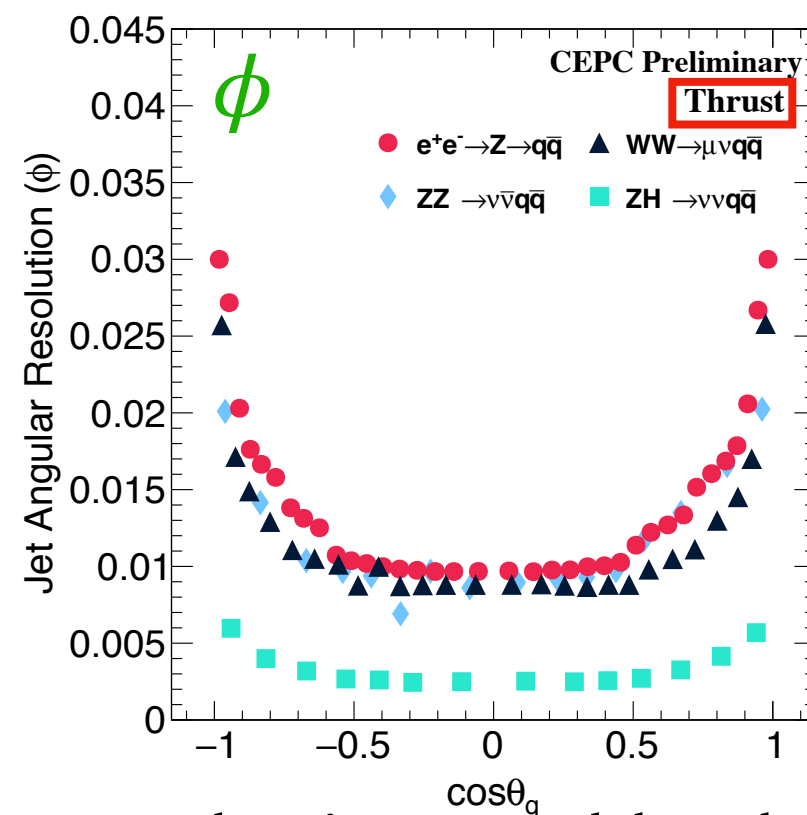
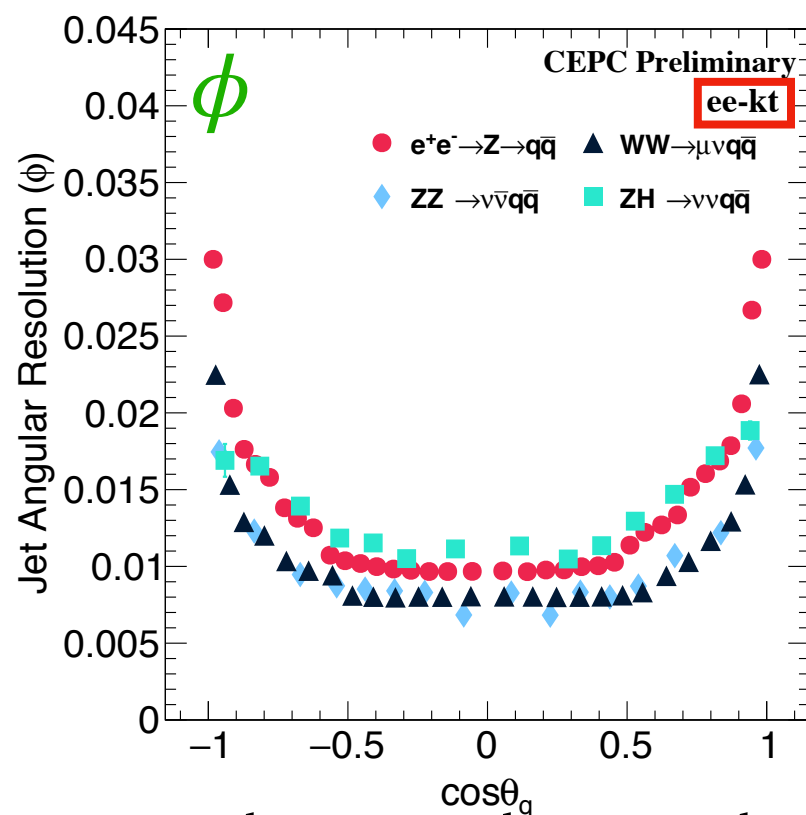
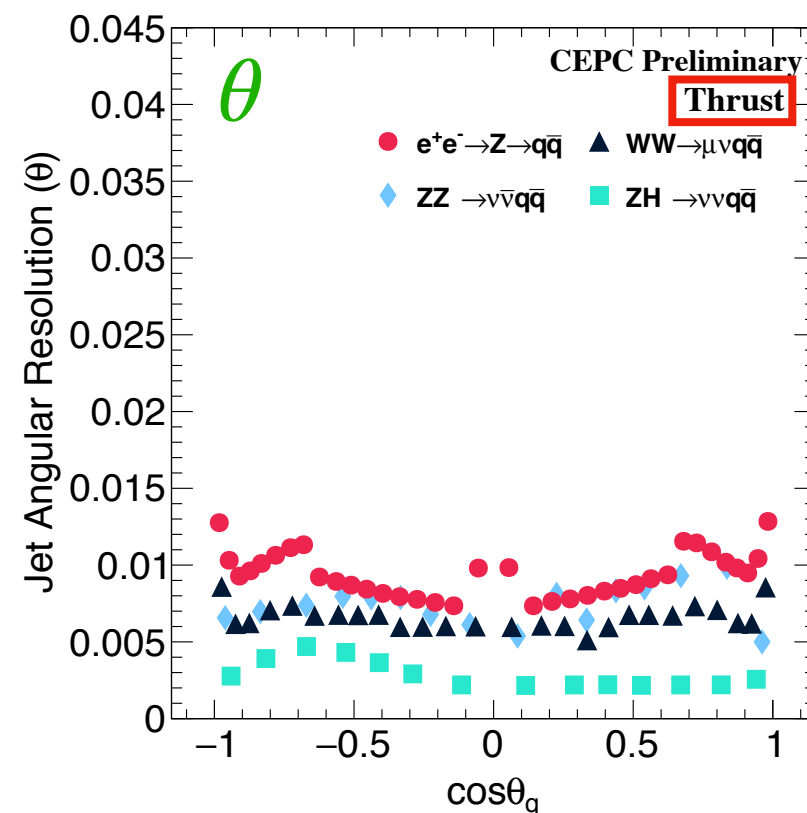
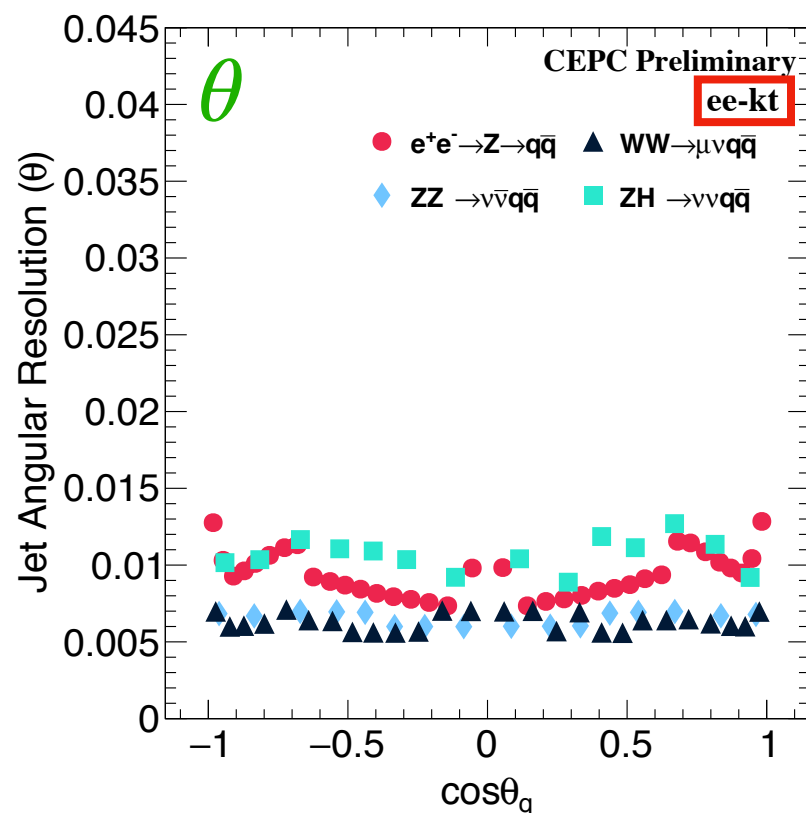
Improved **20%**
w.r.t ee-kt



Improved **40%**
w.r.t ee-kt

■ Improvement maybe came from boosting the system back to the rest frame with the neutrons' information.

BM3: JAR (ee-kt—Thrust)



■ Both of jet θ and ϕ angular resolution are also improved by thrust method,
20%.

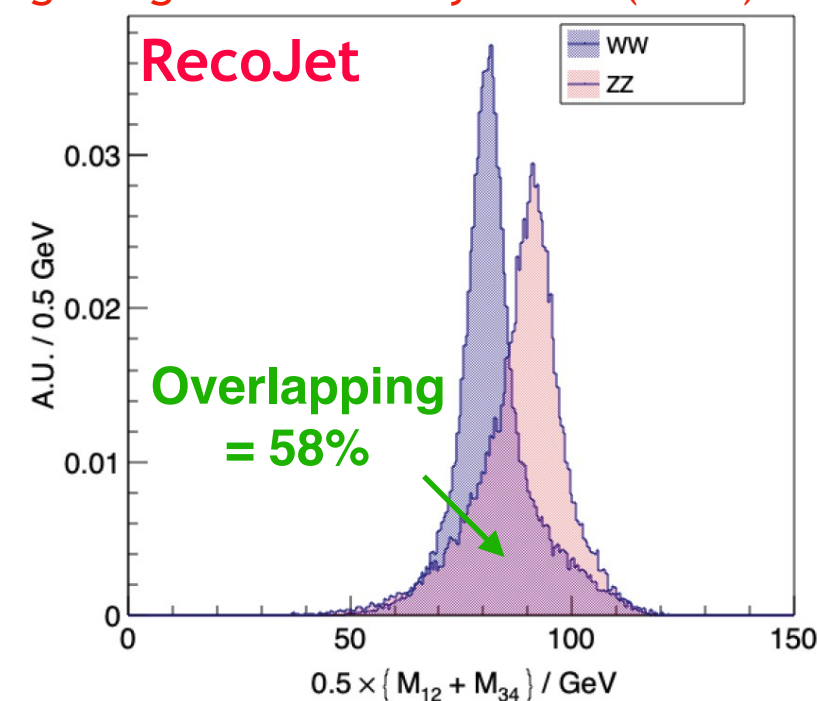
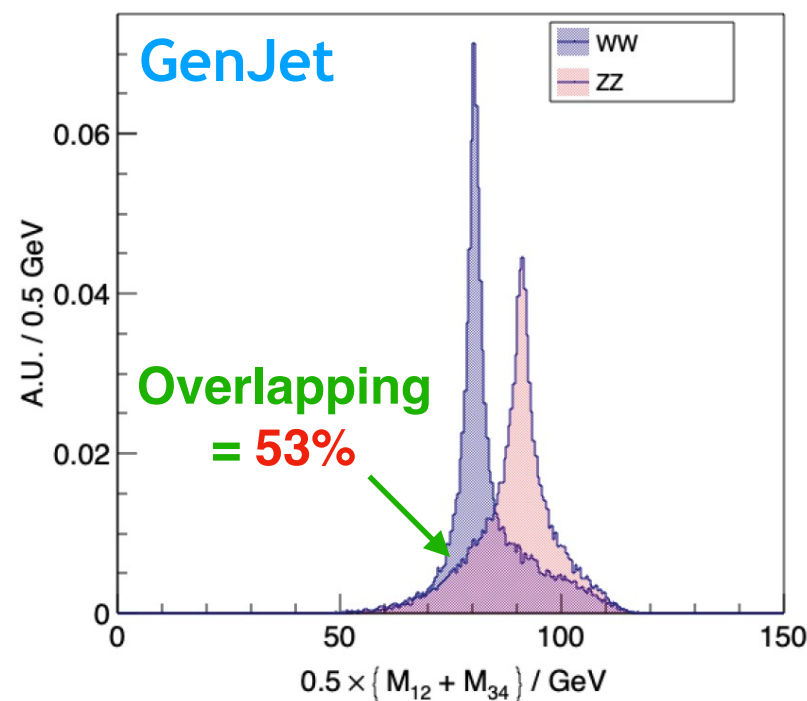
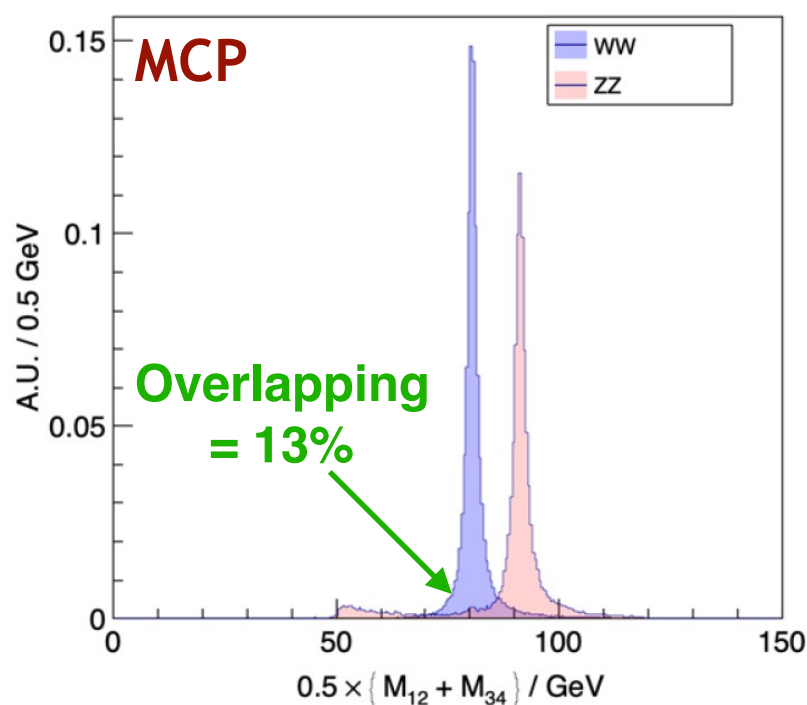
# of jets	Probability	1/3 of ZH events
0	2.44%	<ul style="list-style-type: none"> Major SM Higgs decay modes. 1 color singlet could be identified. (Single Z or Higgs boson)
2	29.73%	
4	59.58%	2/3 of ZH events <ul style="list-style-type: none"> Dominance statistic of $ZH \rightarrow q\bar{q}q\bar{q}$. Major uncertainty is on wrong jet pairing. (Potential huge impact)
6	8.23%	

BM4: Separation of WW , ZZ , and ZH decay to $qqqq$ final state

- 67% (4 + 6 jets) needs **dedicated color-singlet identification**: grouping the hadronic final-state particles into color-singlets (Z, W, H, γ^*). Can be done via jet clustering and pairing.
- **Jet clustering** is also essential for **differential** & EW precision measurements (e.g. TGCs).

BM4: WW & ZZ to 4 Jets Separation

Yong-Feng Zhu—Eur. Phys. J. C (2019) 79:274



- Low energy jet (20-120 GeV)

- Typical multiplicity could be 10^2 .

- GenJet and RecoJet are clustered by ee-kt and paired according to χ^2 .

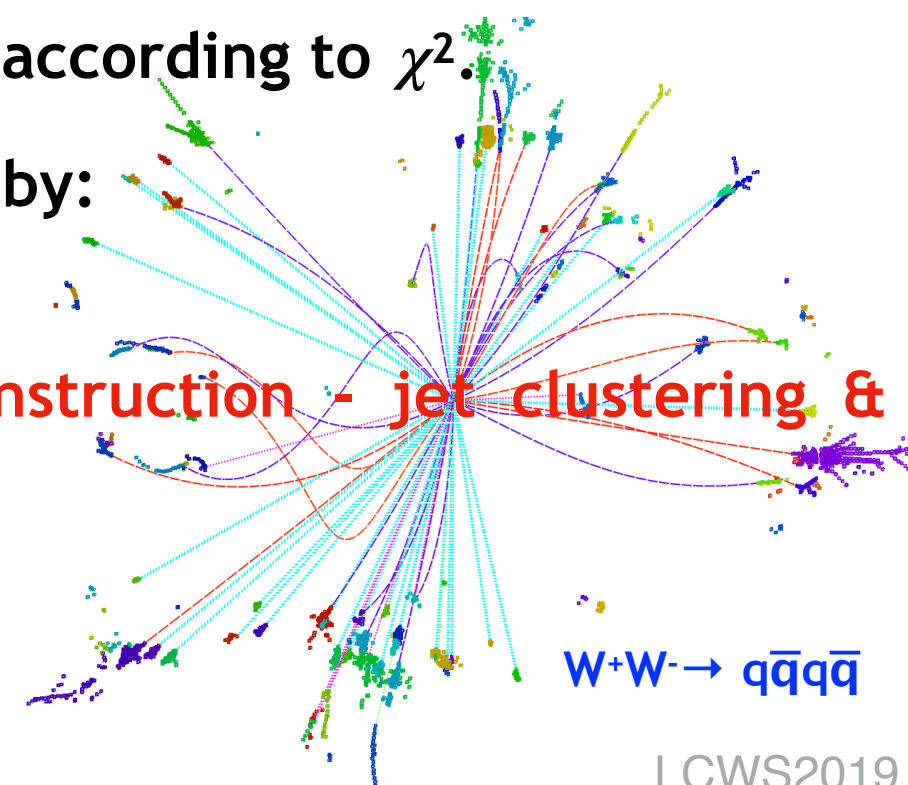
- WW & ZZ to 4 jets final-state separation is determined by:

1. (13%) Intrinsic boson mass/width (10 GeV)
2. (53%) Wrong jet pairing for color singlet reconstruction - jet clustering & pairing.
3. (58%) Detector response

40%

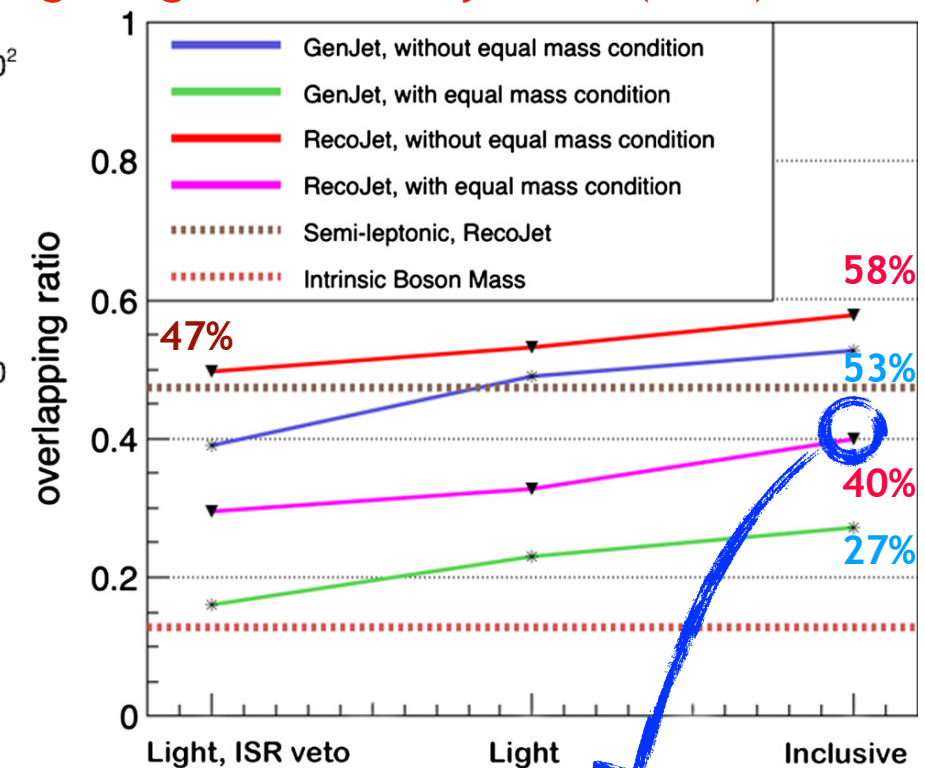
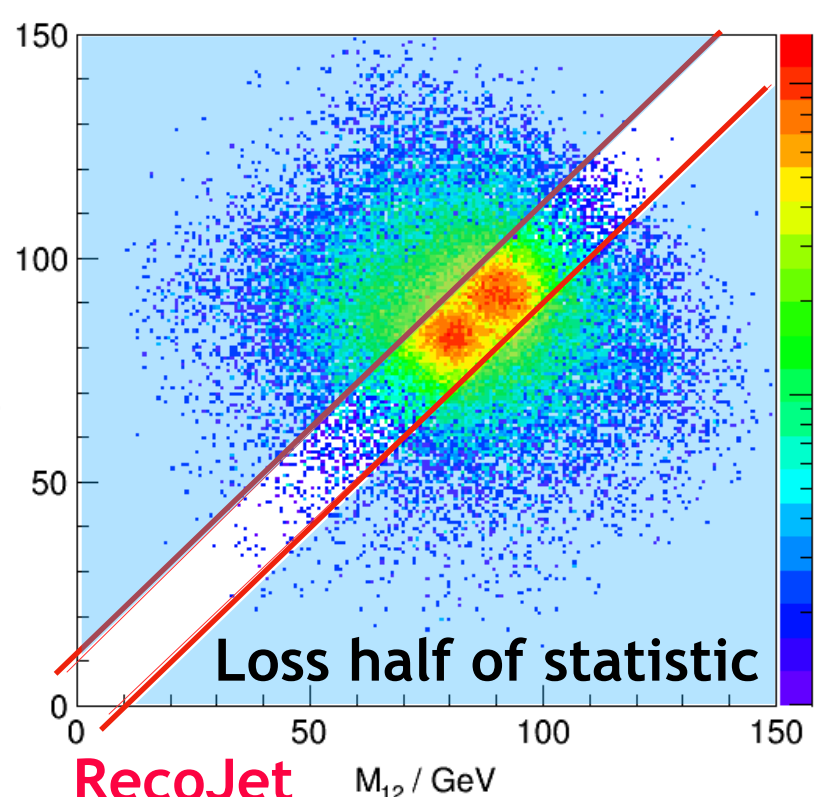
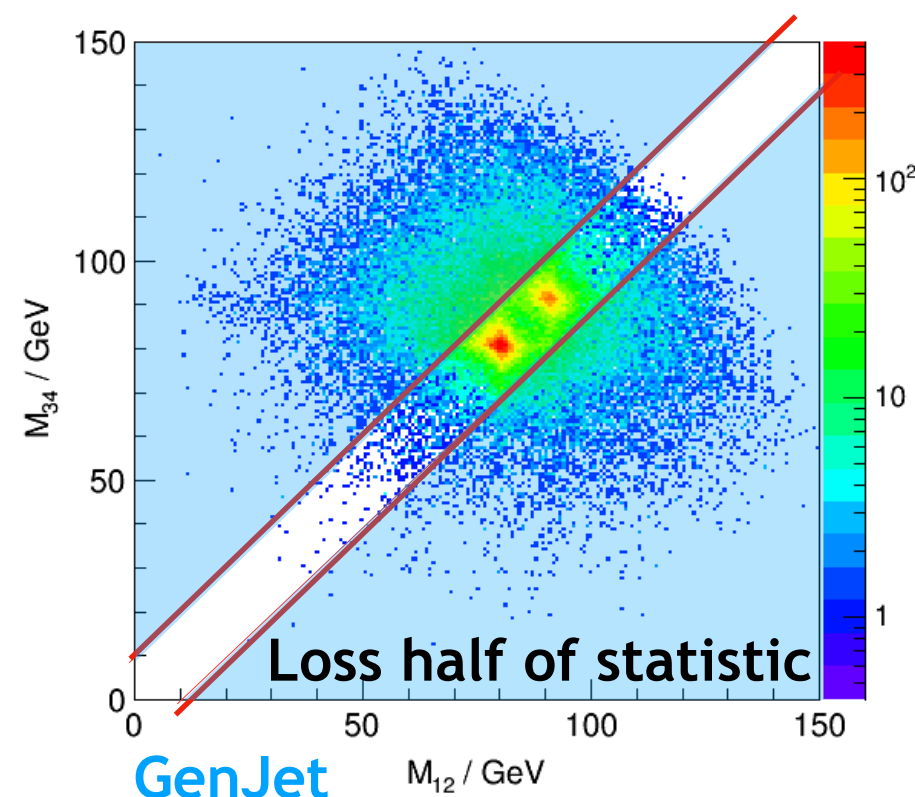
5%

$$\chi^2 = \frac{(M_{12} - M_B)^2 + (M_{34} - M_B)^2}{\sigma_B^2}$$



BM4: WW & ZZ to 4 Jets Separation

Yong-Feng Zhu—Eur. Phys. J. C (2019) 79:274



■ Equal mass requirement: $|M_{12} - M_{34}| < 10$ GeV

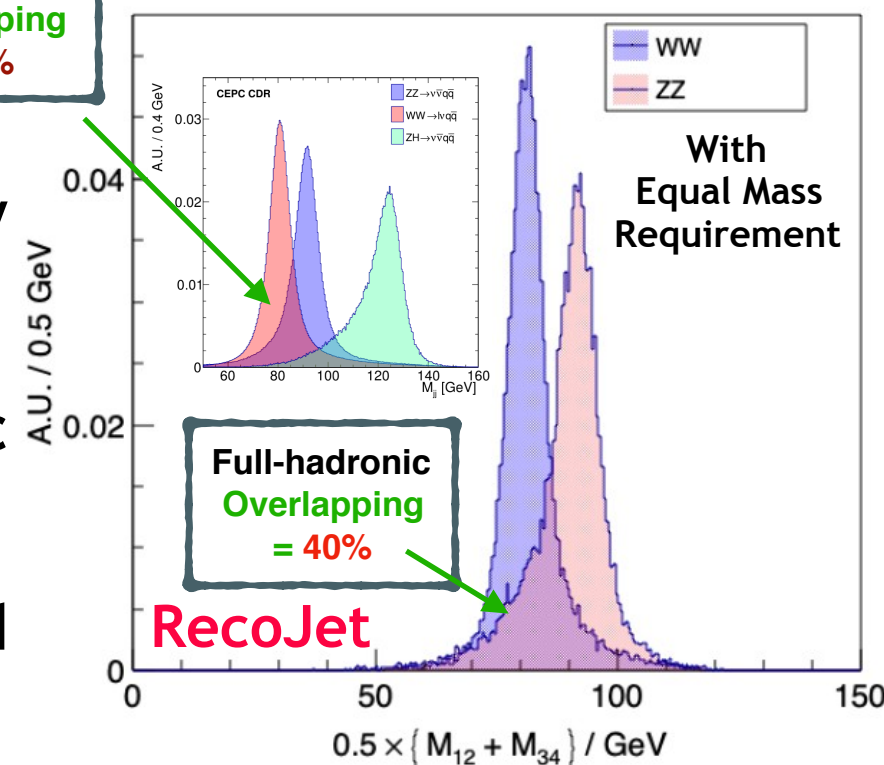
- Cost half of the statistic.
- Overlapping can be reduced from **58%/53%** to **40%/27%** for the **RecoJet/GenJet**.

■ CEPC baseline could separate WW & ZZ with full hadronic final-state.

■ Improve from the naive jet clustering & pairing and control the ISR photon in the event.

■ ZH full hadronic final-state analysis is on the way.

Semi-leptonic
Overlapping
= 47%



Full-hadronic
Overlapping
= 40%

■ Jets are crucial for the CEPC Higgs physics

- 97% of ZH events evolve jets
- 1/3 only come from single Z or Higgs boson.
- 2/3 has more than one boson (e.g. $ZH \rightarrow q\bar{q}q\bar{q}$) - need color singlet identification algorithm.

I. **BMR < 4% is critical. Achieved at the CEPC baseline (3.8%)**

- * W, Z, Higgs boson can be efficiently separated at both semi-leptonic & full hadronic.
- * By Z-boson di-jet recoil mass to distinguish the ZH from ZZ process.

II. **2 jets final-state could be identified with $efficiency \times purity = 88.4\%$.**

- * Could be clustered by dedicated jet clustering algorithm, thrust.

III. **Single Jet — JER ~ 3-5% & JAR ~ 1%.**

- * Thrust clustering method is recommended for two jets final-state. It could improve the JER 20%, 40% on tail (RMS), and JAR 20%.

IV. **Need a better color singlet identification algorithm.**

- * Wrong jet pairing is the dominant effect to induce overlapping in full hadronic WW-ZZ separation.
- * Equal mass requirements: Reduce the overlapping to be better than semi-leptonic, but very costly.
- * Other physical impact is significant: ISR photon etc.

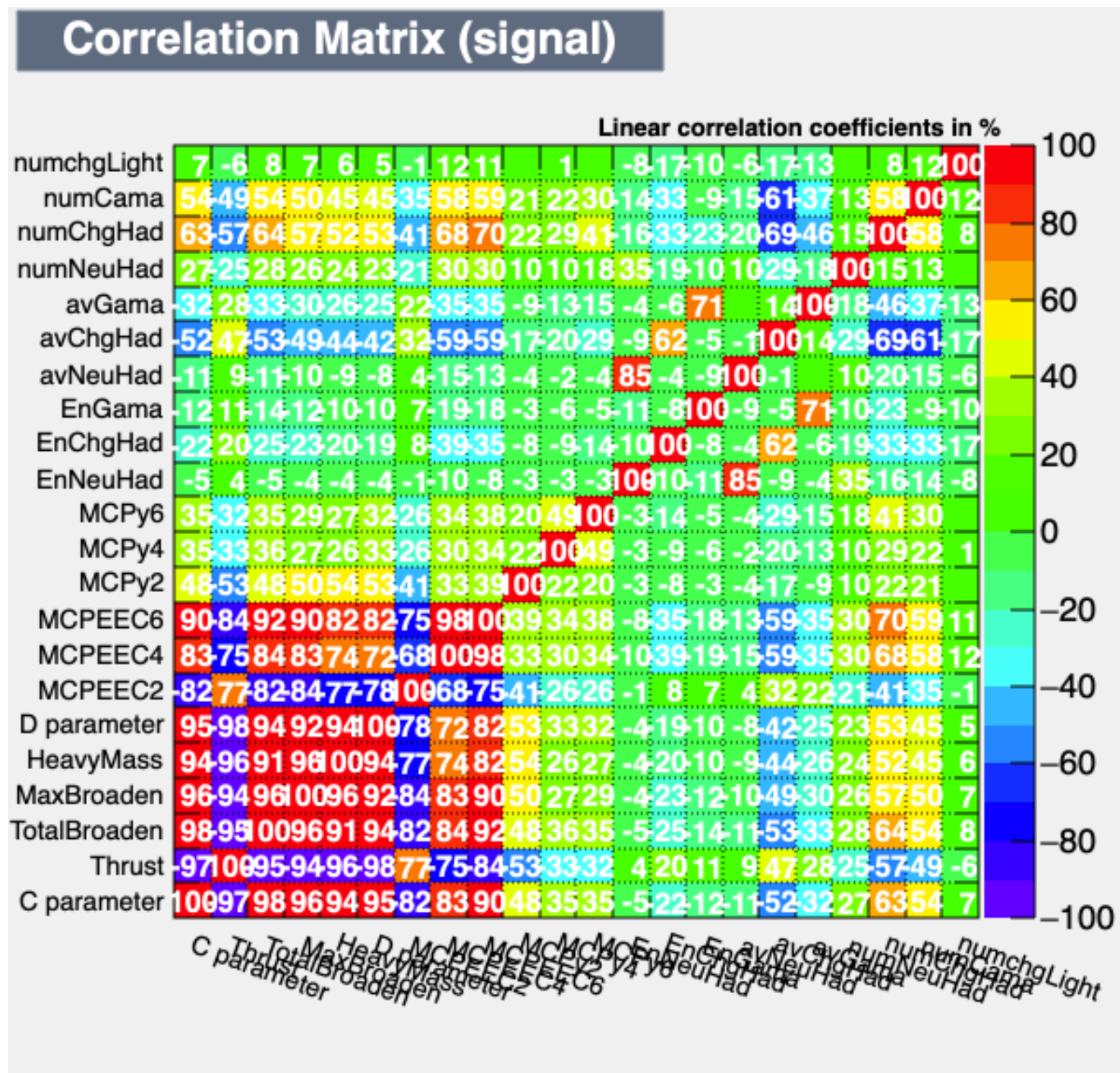
Thank for your attention



BM2: Number of Jet Identification

20 Variables

# of charge lepton	EEC 6
# of γ	EEC 4
# of charge hadron	EEC 2
# of neutro hadron	C parameter
\bar{E}_γ	D parameter
$\bar{E}_{\text{Charge hadron}}$	Heavy Mass
$\bar{E}_{\text{Neutro hadron}}$	Max Broaden
E_γ	Total Broaden
$E_{\text{Charge hadron}}$	Thrust
$E_{\text{Neutro hadron}}$	y_{23}, y_{45}, y_{67}



- Event-shape variables basic multi-variable analysis to separate 2, 4, and 6 jets final-state.

Yong-Feng Zhu

Event-shape Variables

Heavy Jet Mass

$$M_1^2 = \frac{1}{(\sqrt{s})^2} \left(\sum_i^N P_i \right)^2$$

$$M_2^2 = \frac{1}{(\sqrt{s})^2} \left(\sum_i^N P_i \right)^2$$

Jet Broadening

$$B_1 = \frac{1}{2 \sum_{j=1}^N |P_j|} \sum_{i=1}^N |P_i \times n_T|, (P_i \times n_T) > 0$$

$$B_2 = \frac{1}{2 \sum_{j=1}^N |P_j|} \sum_{i=1}^N |P_i \times n_T|, (P_i \times n_T) < 0$$

Jet Transition variable, y_{23} , y_{45} , y_{67}

ee-kt jet clustering algorithm

$$d_{ij} = 2\min(E_i^2, E_j^2)(1 - \cos\theta_{ij})$$

C and D Parameter

$$L^{ab} = \frac{1}{\sum_{j=1}^N |P_j|} \sum_{i=1}^N \frac{P_i^a P_i^b}{|P_i|}$$

$$C = 3(\lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3)$$

$$D = 27\lambda_1\lambda_2\lambda_3$$

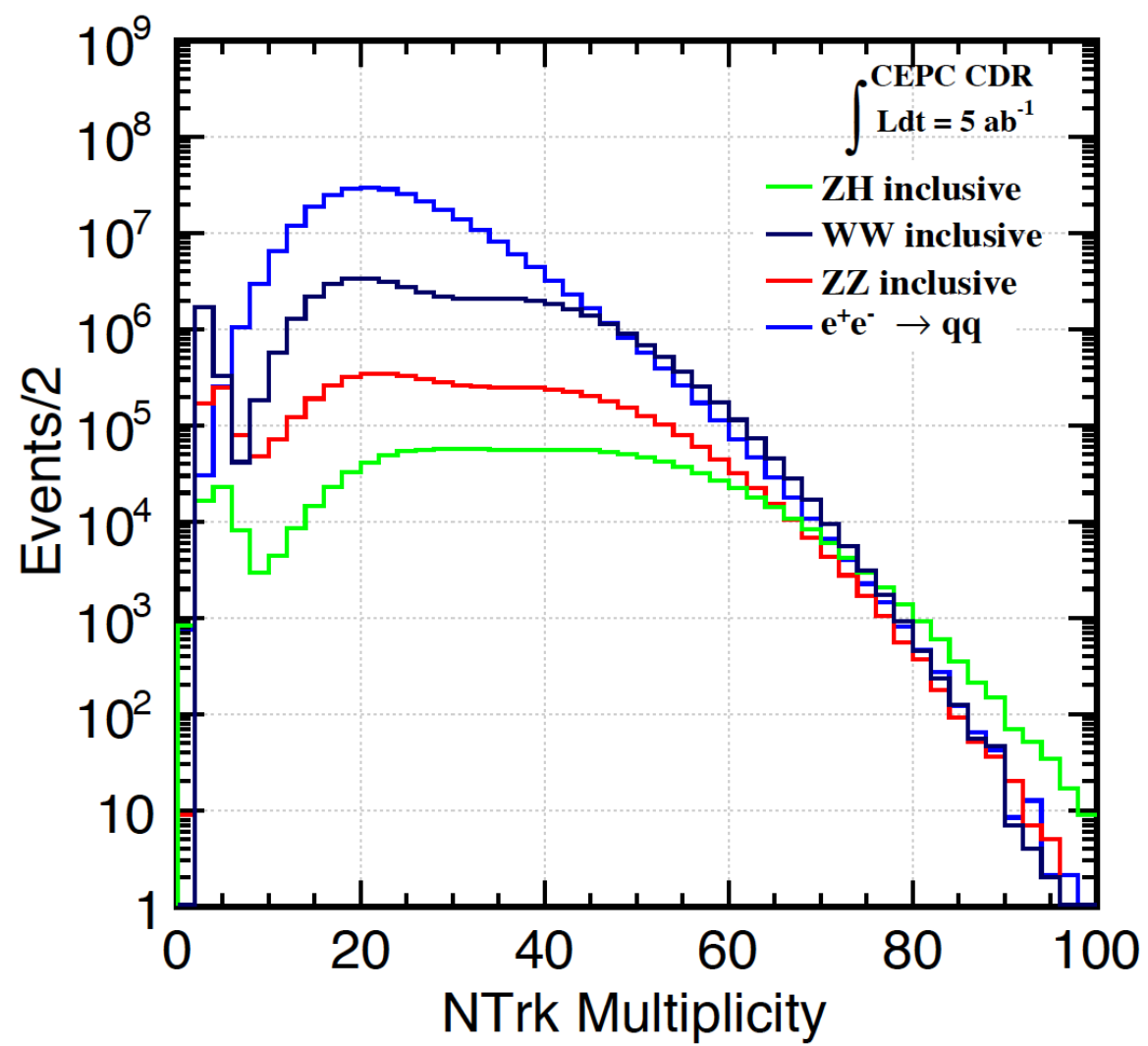
Energy-Energy Correlation

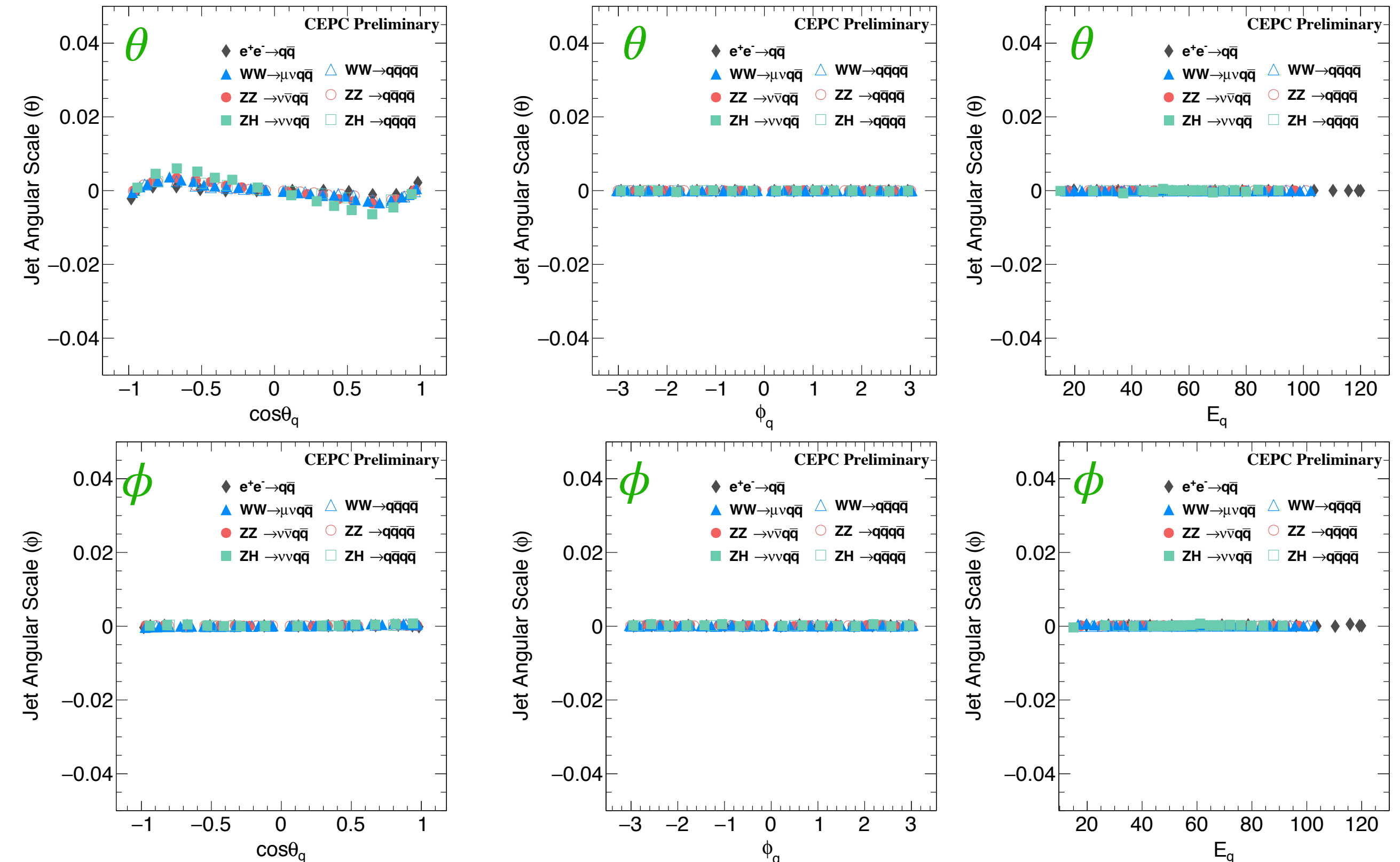
$$EEC = \frac{1}{\sigma_{tot}} \sum_{ij} \int d\sigma \frac{E_i E_j}{Q^2} \delta(\cos\chi - \cos\theta_{ij})$$

$$likelihood = \frac{\sum (P1_i) \times P2_i}{\sqrt{\sum (P1_i \times P2_i) \times \sum (P2_i \times P2_i)}}$$

Double-sided Crystal Ball

$$f(x|\alpha_1, \alpha_2, n_1, n_2, \bar{x}, \sigma) = \begin{cases} \left(\frac{n_1}{|\alpha_1|}\right)^{n_1} e^{-\frac{|\alpha_1|^2}{2}} \left(\frac{n_1}{|\alpha_1|} - |\alpha_1| - \frac{x - \bar{x}}{\sigma}\right)^{-n_1} & \frac{x - \bar{x}}{\sigma} < -\alpha_1 \\ e^{-\frac{1}{2}\left(\frac{x - \bar{x}}{\sigma}\right)^2} & -\alpha_1 < \frac{x - \bar{x}}{\sigma} < \alpha_2 \\ \left(\frac{n_2}{|\alpha_2|}\right)^{n_2} e^{-\frac{|\alpha_2|^2}{2}} \left(\frac{n_2}{|\alpha_2|} - |\alpha_2| - \frac{x + \bar{x}}{\sigma}\right)^{-n_2} & \alpha_2 < \frac{x - \bar{x}}{\sigma} \end{cases}$$

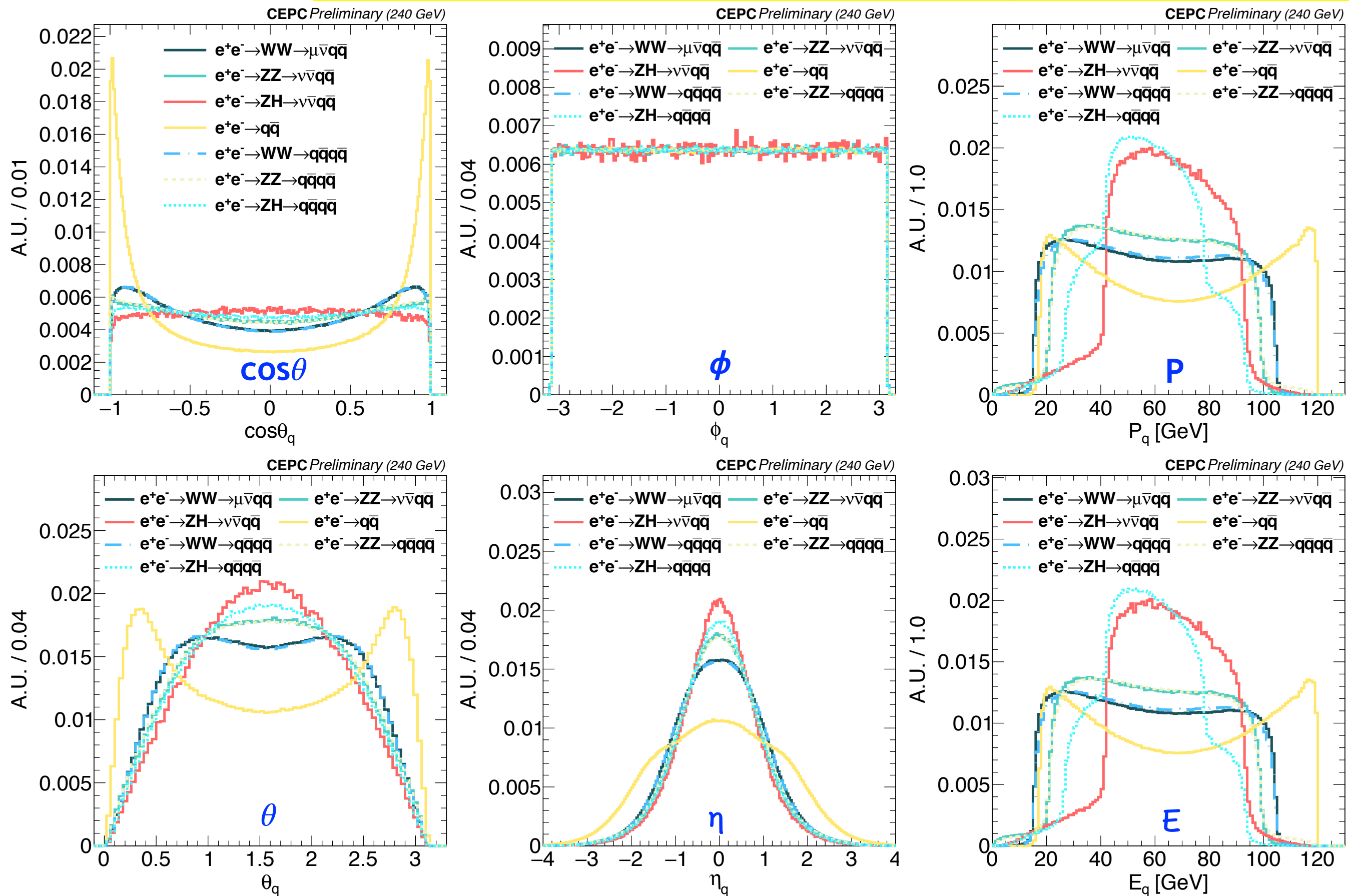




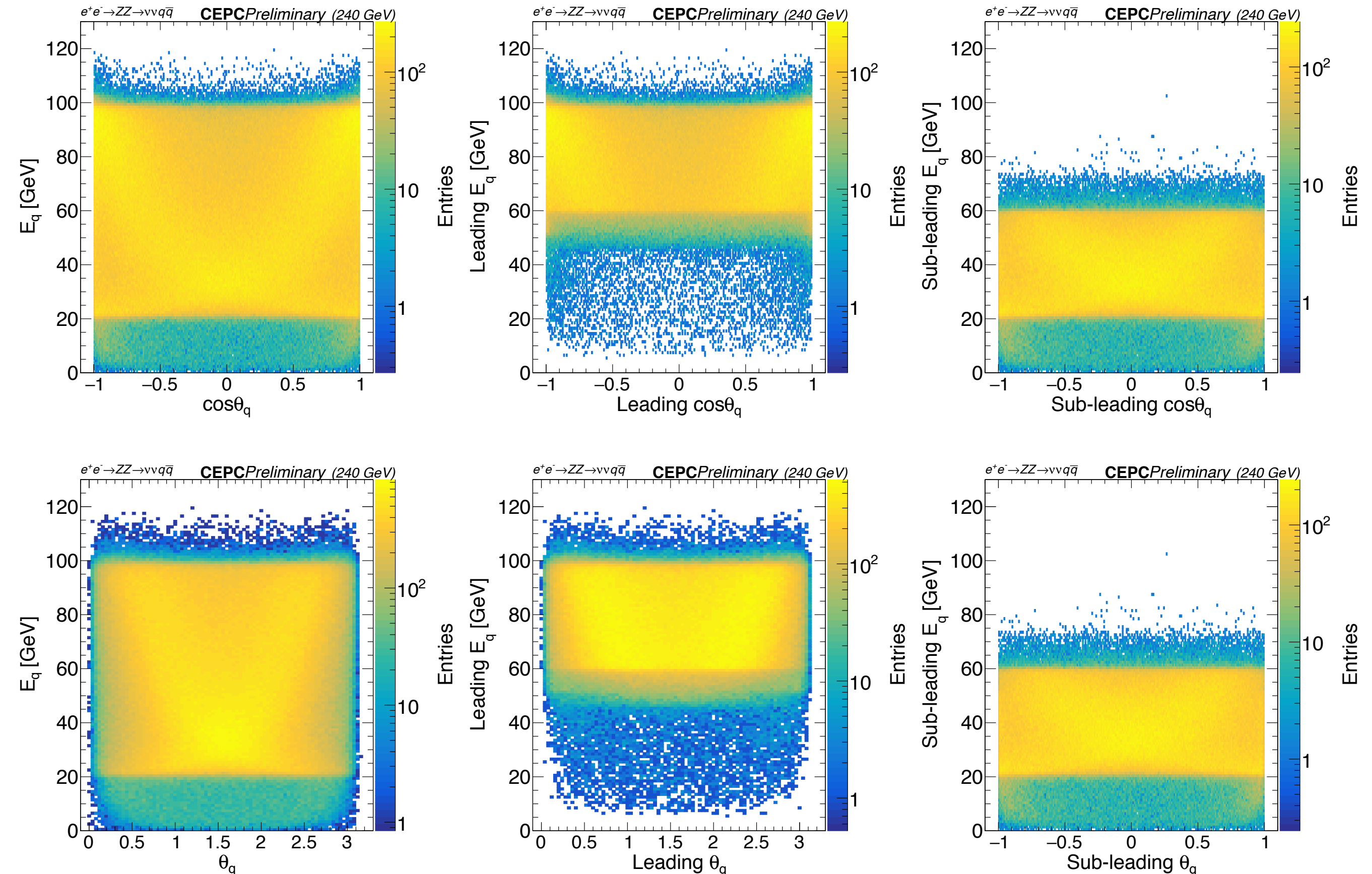
■ JAR is around **1%** in barrel region; JAS is **independent** of ϕ and energy.

■ The difference between 2 and 4 jets final-state is controlled within **1%** level.

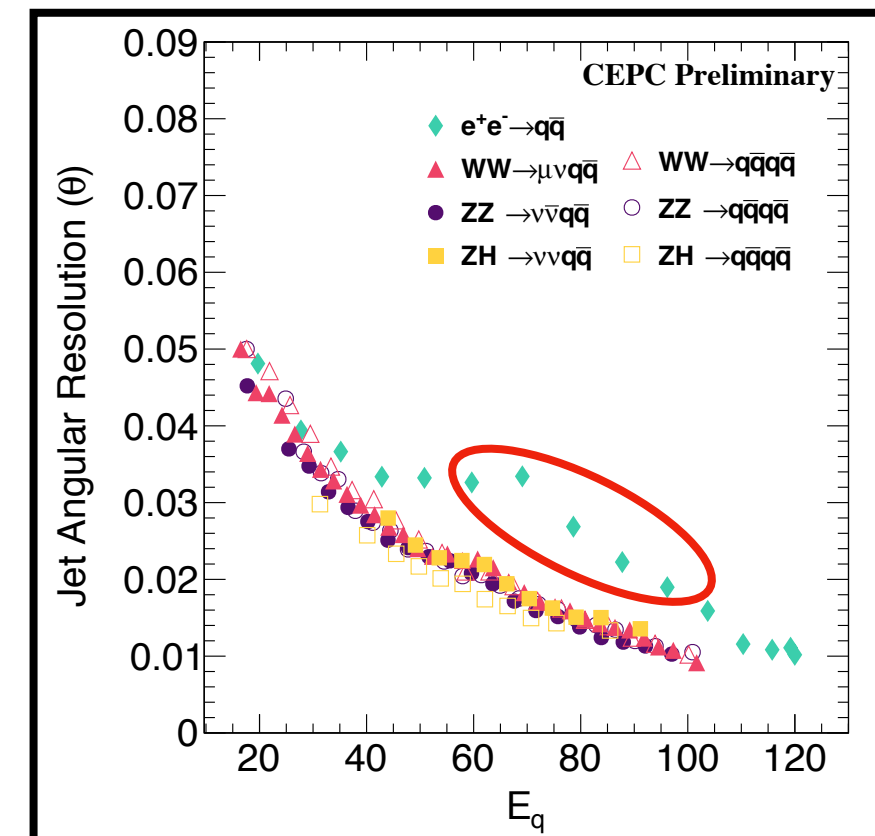
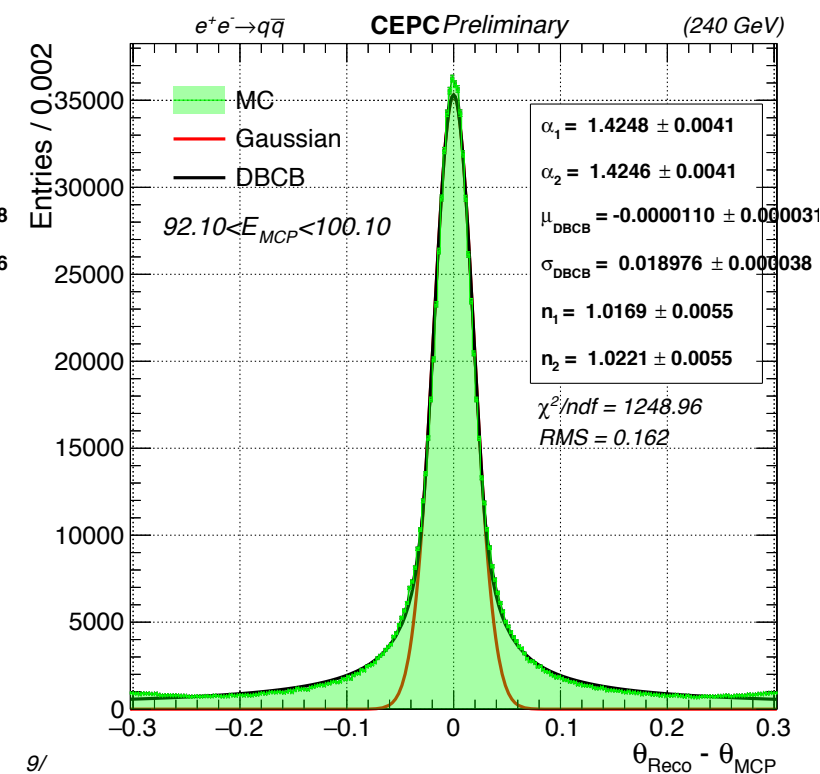
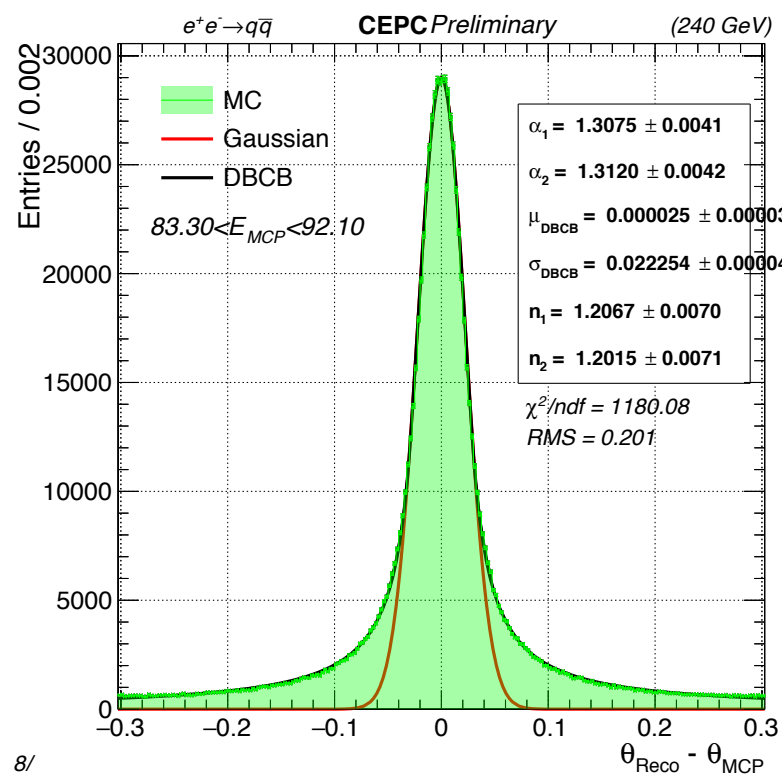
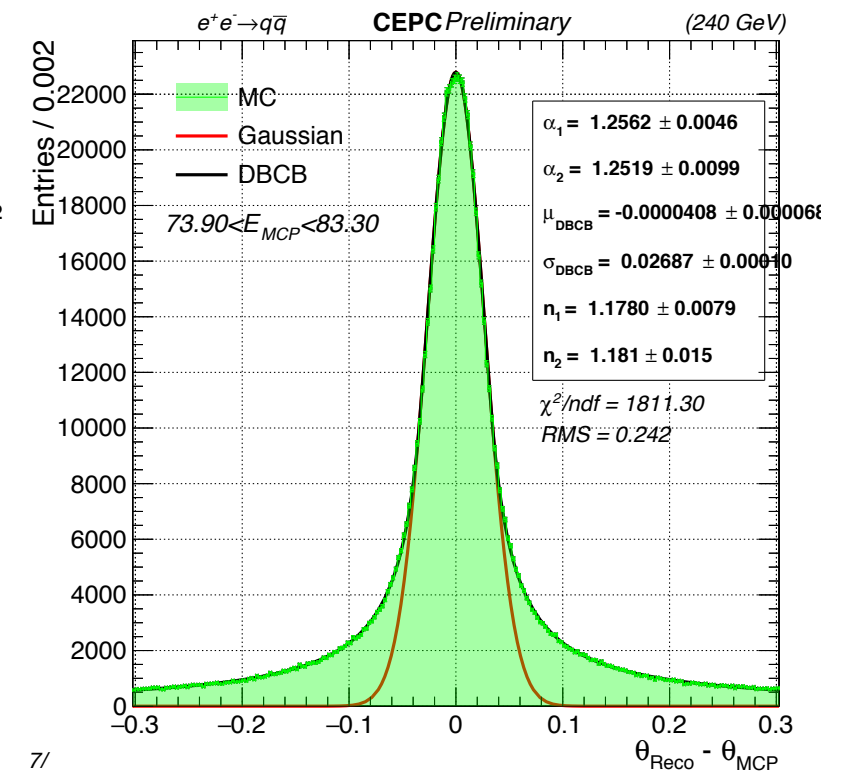
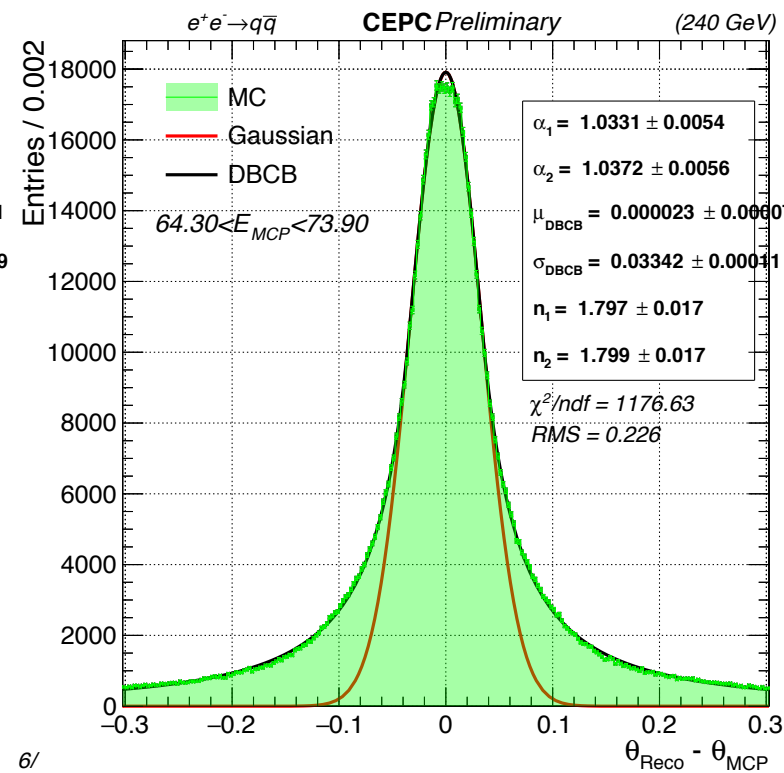
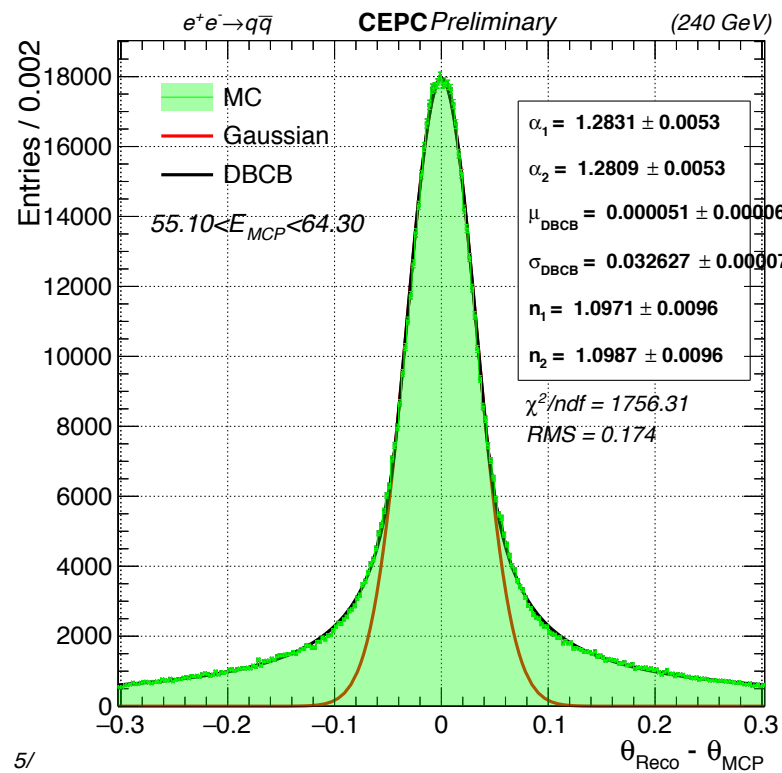
Kinematic Summary Plots (Parton level)



E as the function of the polar angle

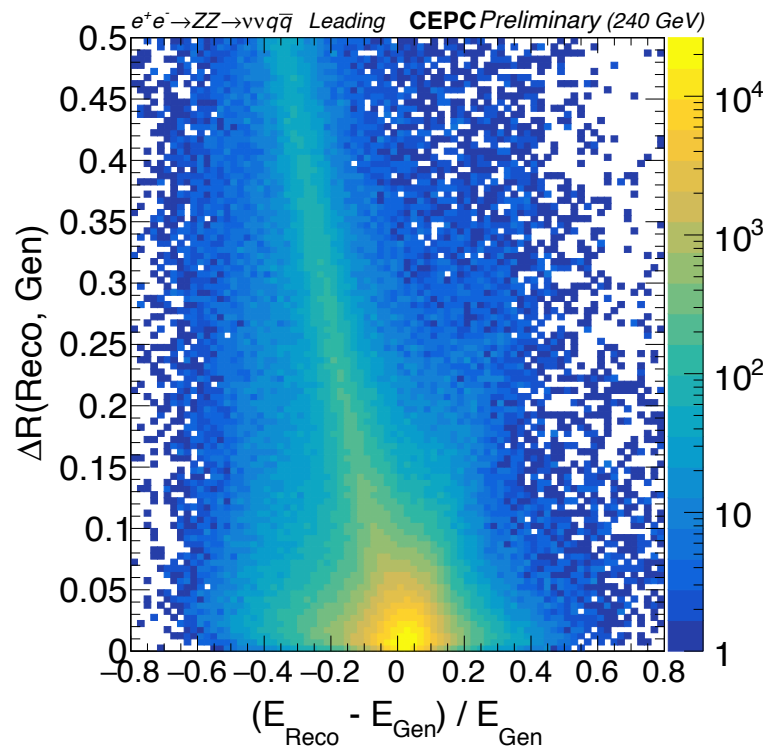


JAR(Reco-MCP)

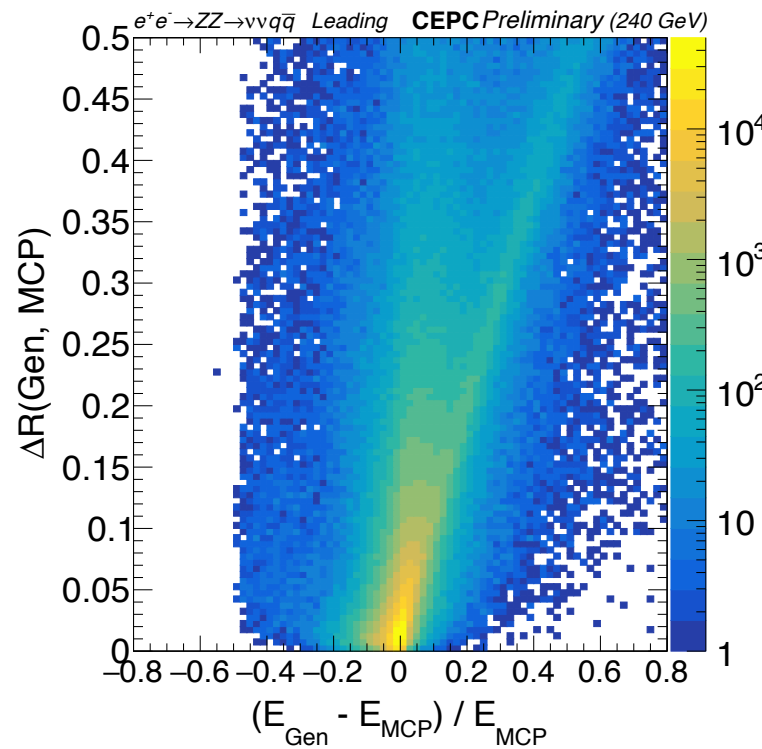


ΔR as the function of ΔE

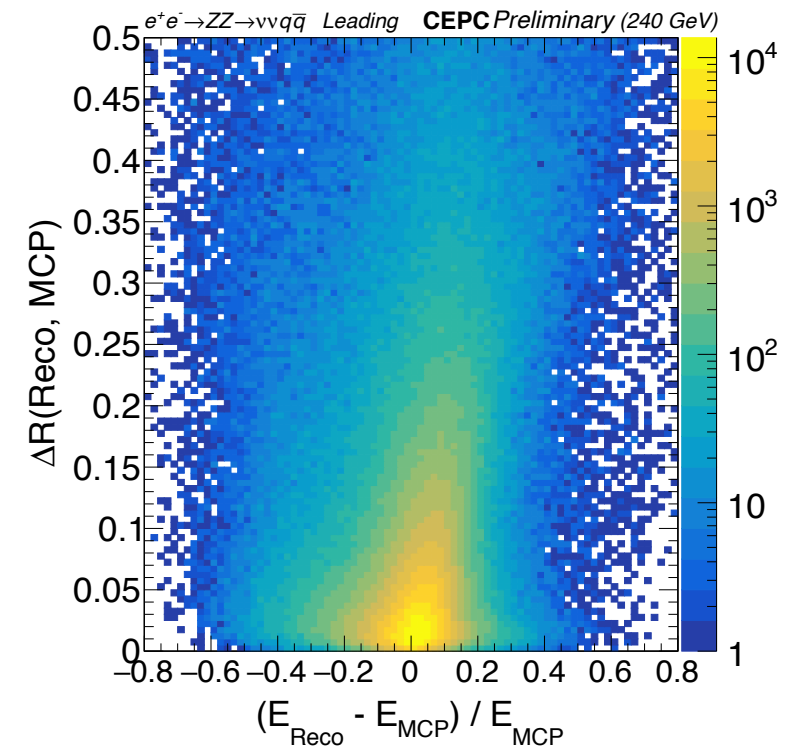
Leading, Gen-MCP



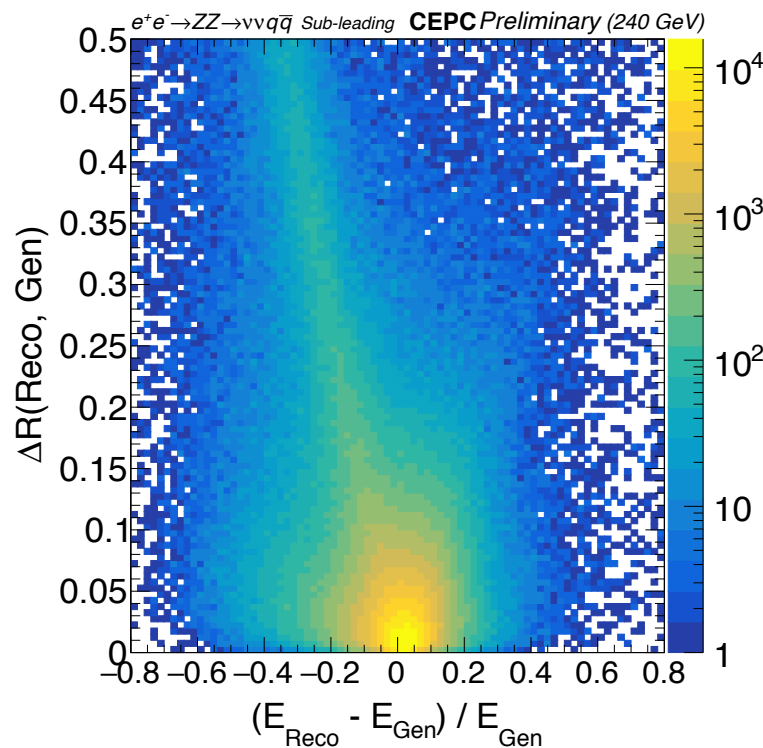
Leading, Reco-Gen



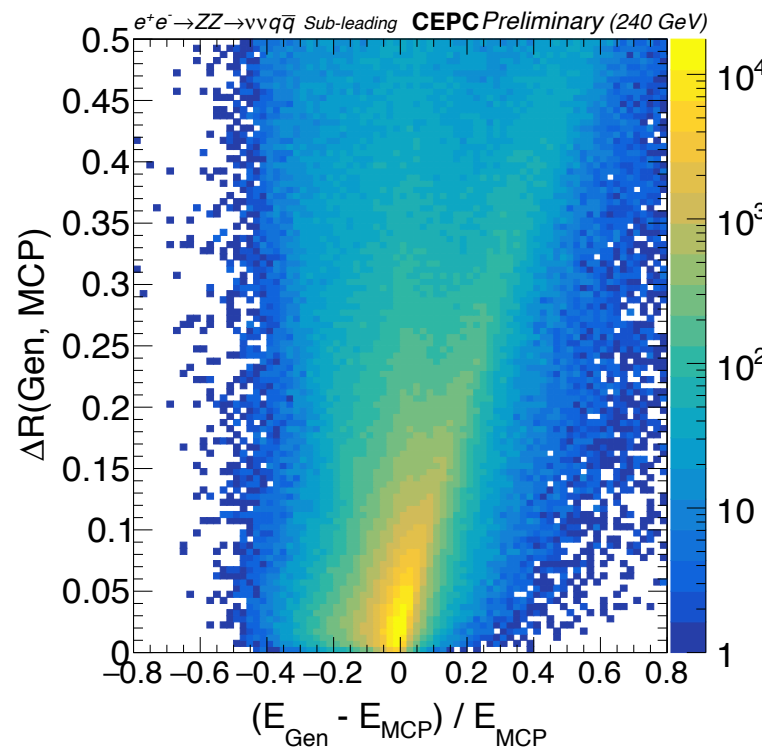
Leading, Reco-MCP



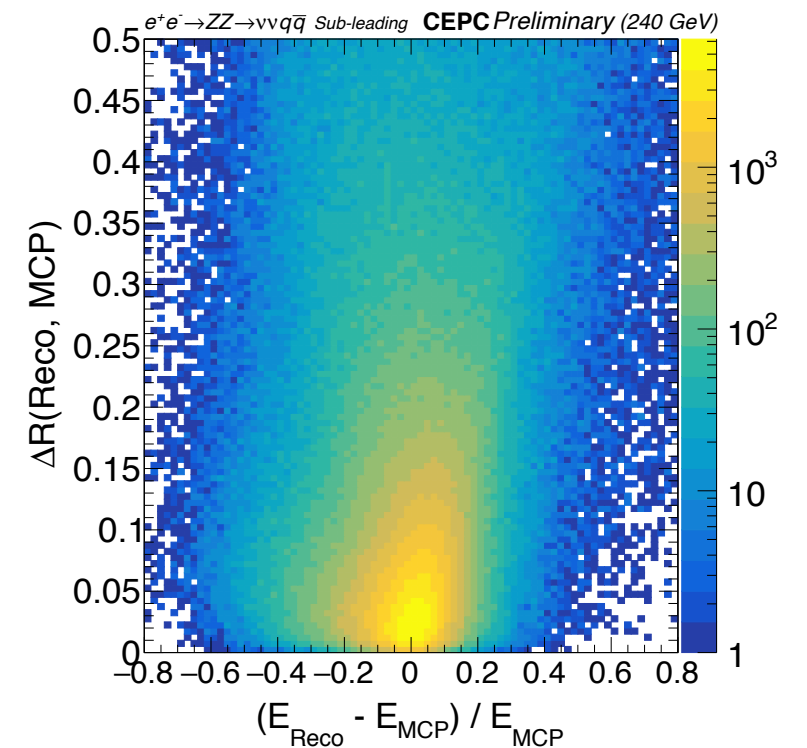
Sub-leading, Gen-MCP



Sub-leading, Reco-Gen

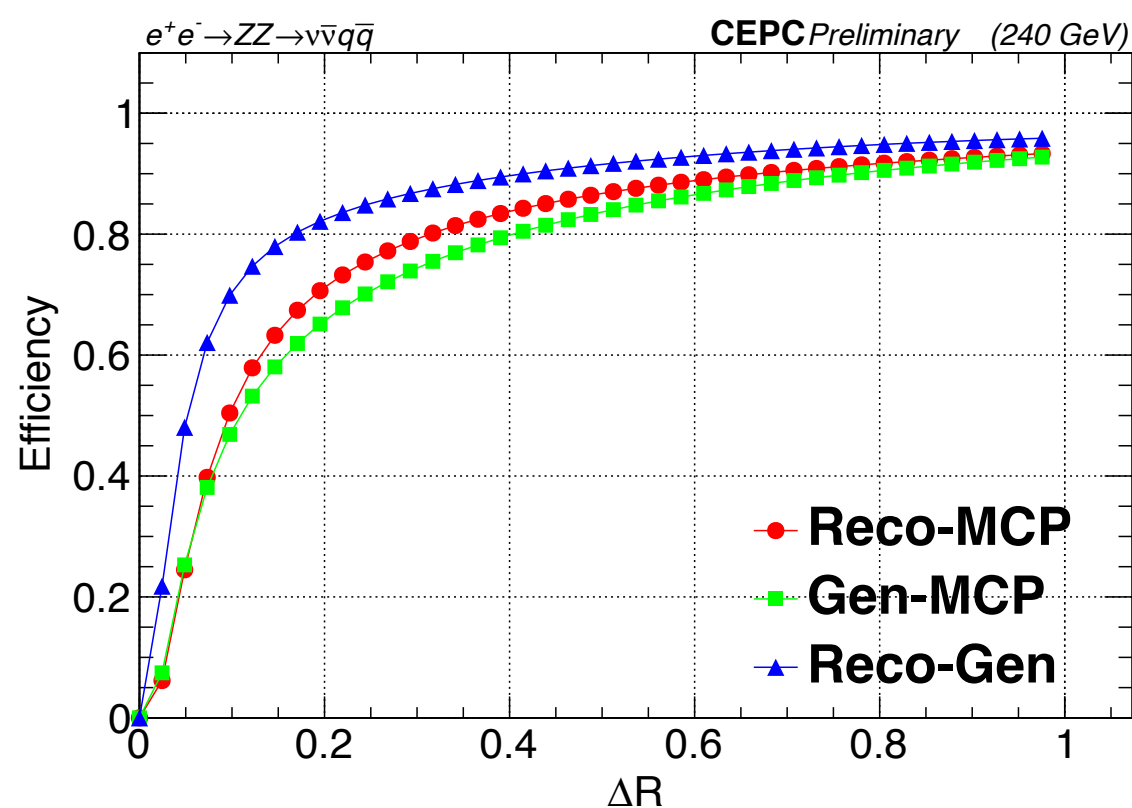
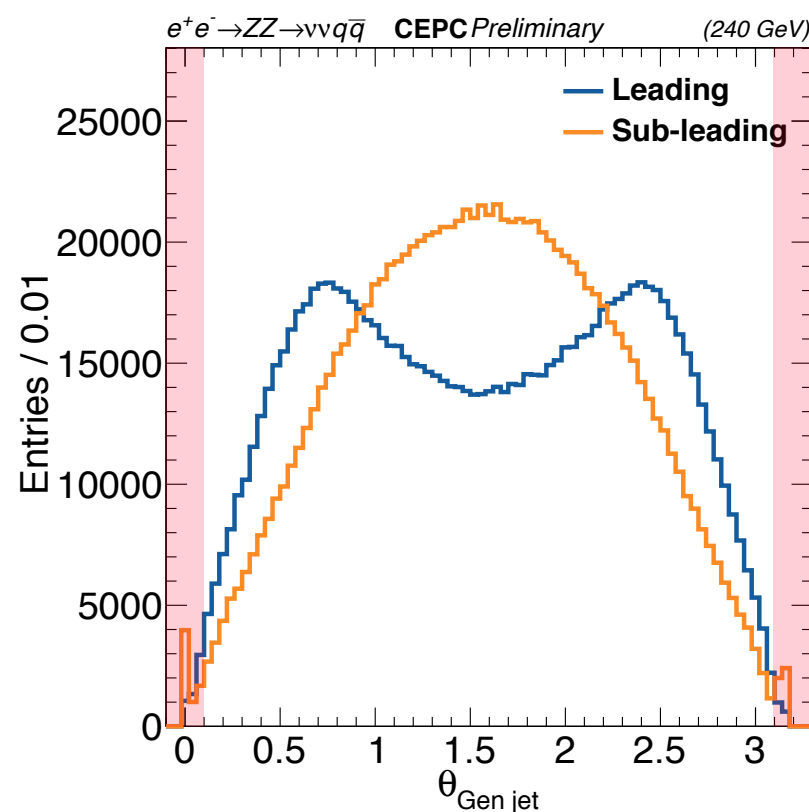


Sub-leading, Reco-MCP



The jet clustering brings a significant uncertainty.

Items	(Reco-Gen)	(Gen-MCP)
$\theta_{\text{Gen jet}} > 0.1 \ \& \ \theta_{\text{Gen jet}} < 3.1$	✓	✓
$\Delta R(\text{Reco-MCP}) < 0.1$	✓	✗

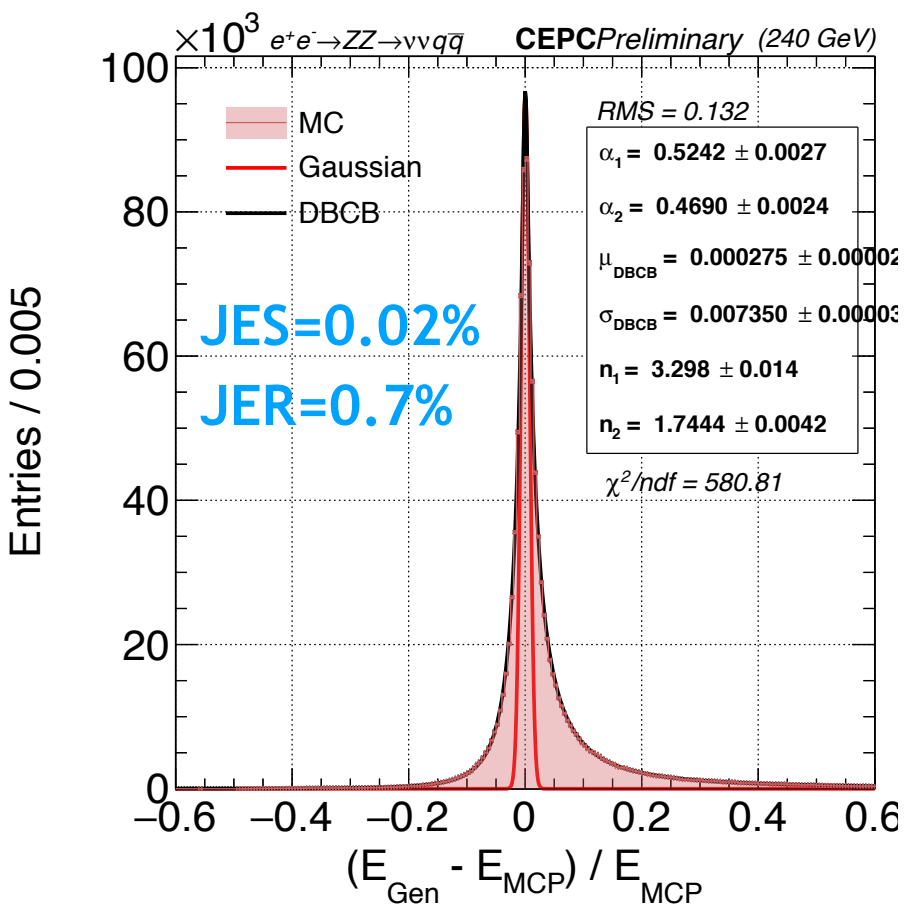


Efficiency =

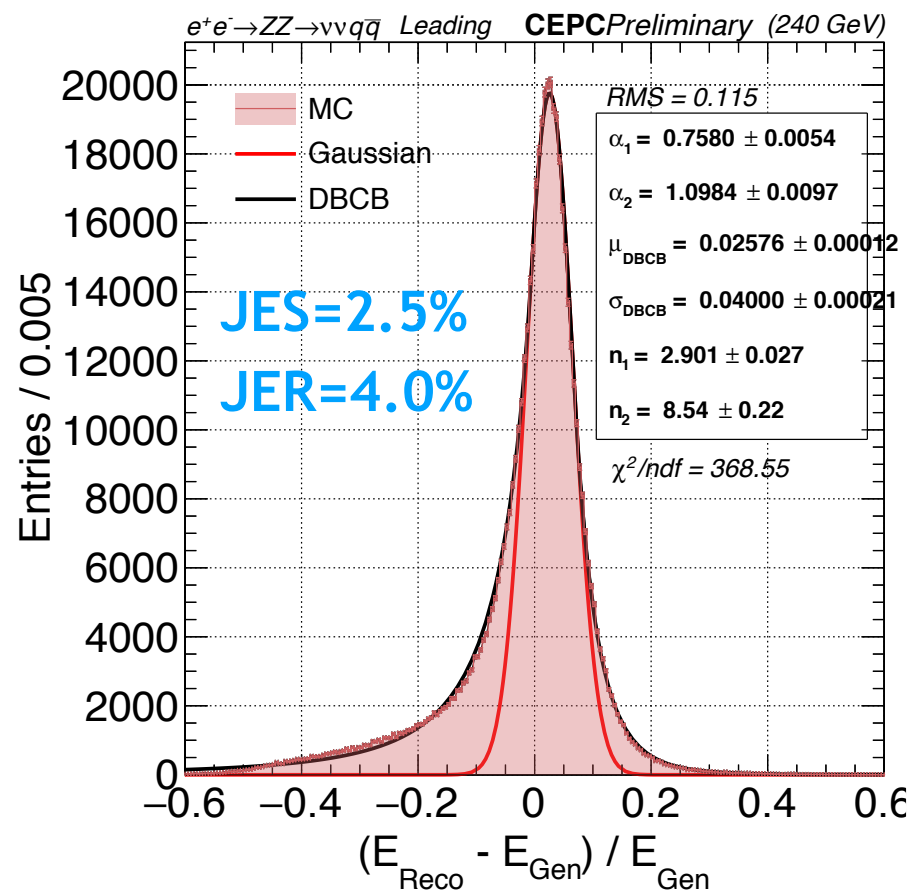
$$\frac{\text{\# of leftover event}}{\text{\# of total event}}$$

$$\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$$

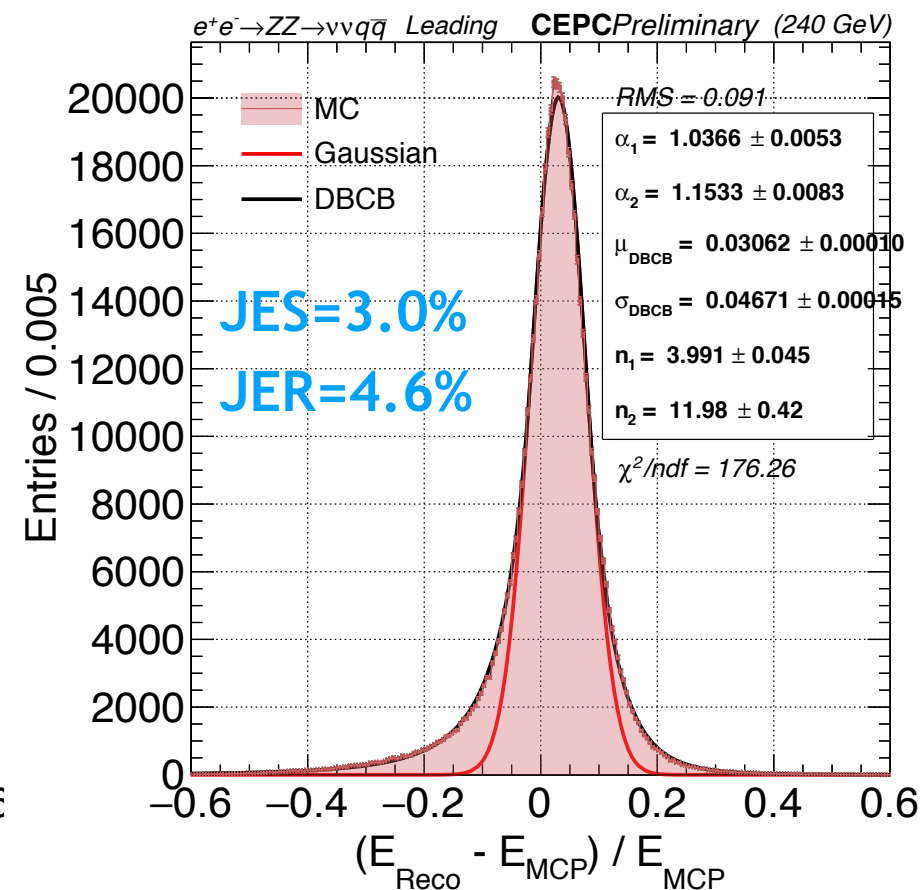
Gen-MCP



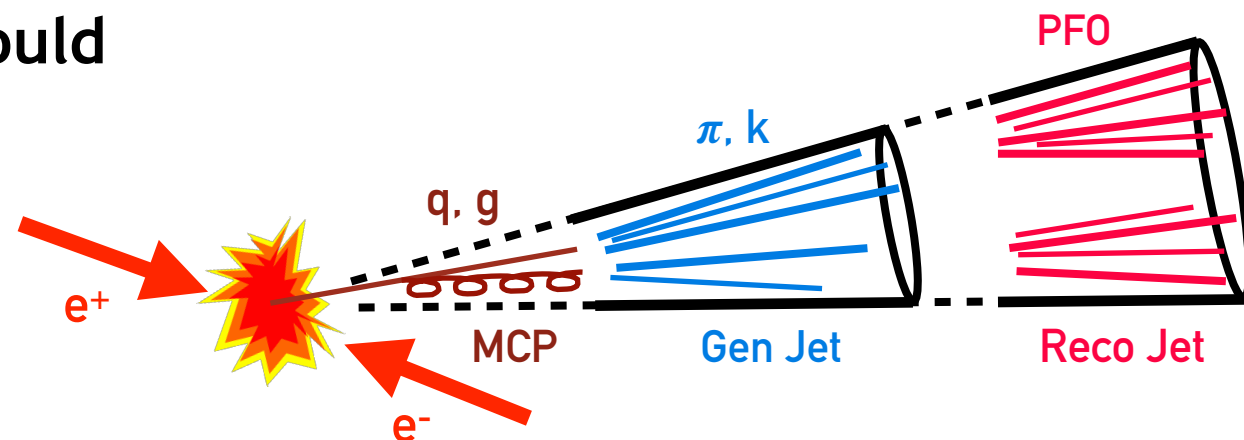
Reco-Gen



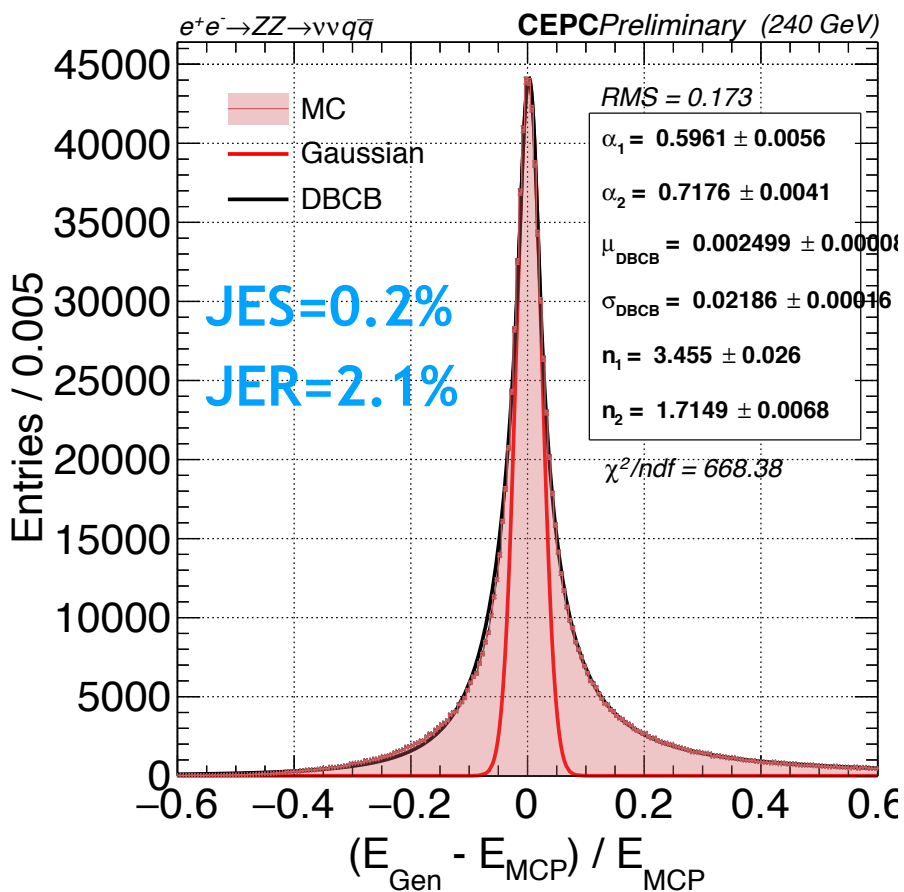
Reco-MCP



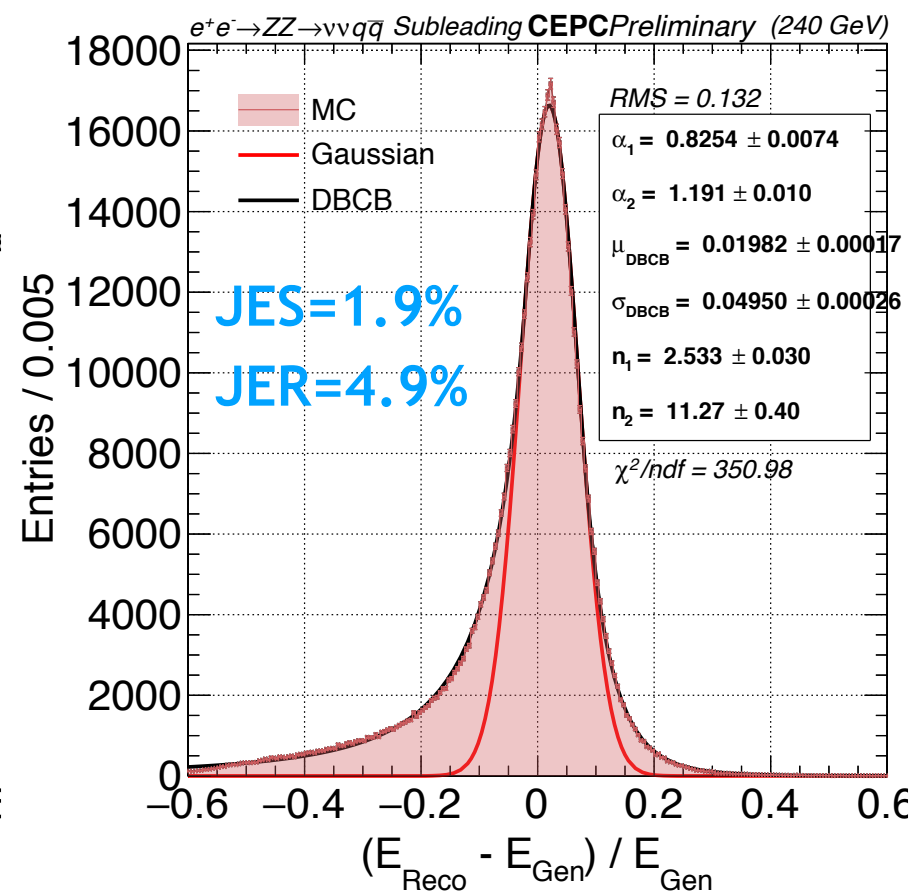
- JER/JES between Reco jet and MCP would combine the effects of two previous stages.



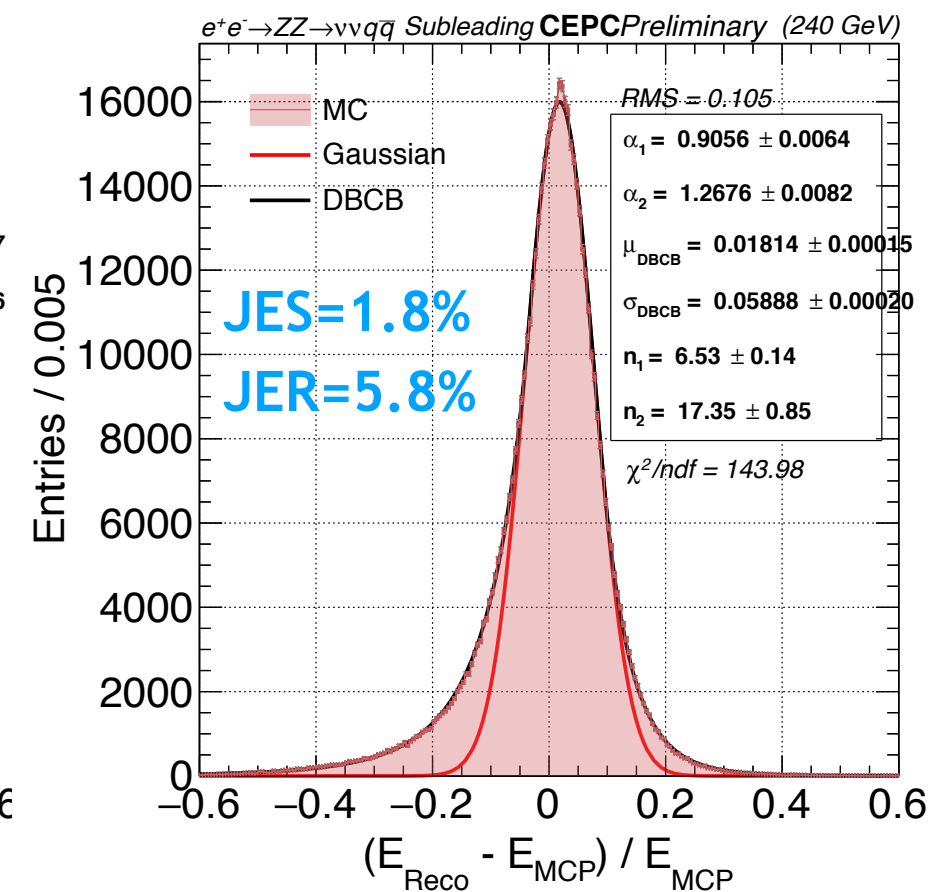
Gen-MCP



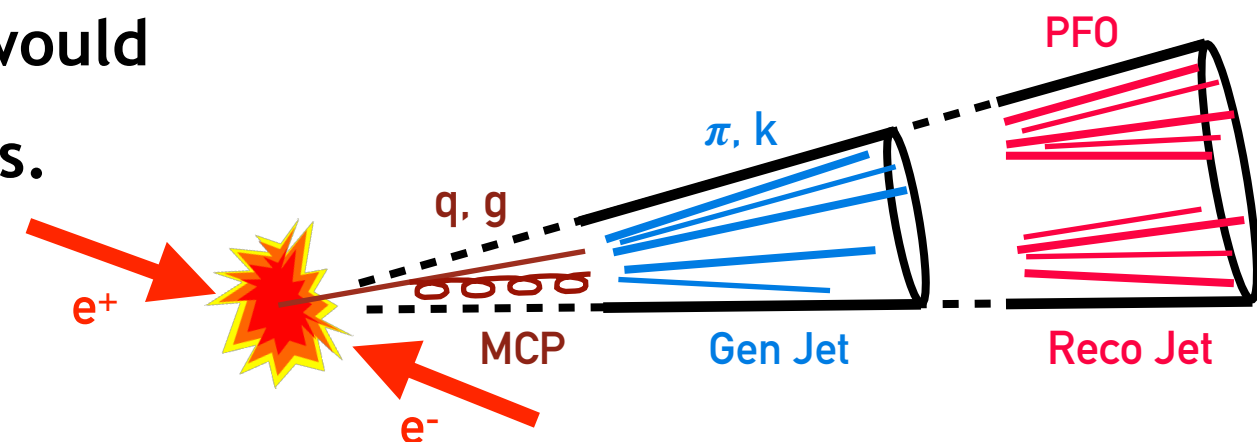
Reco-Gen



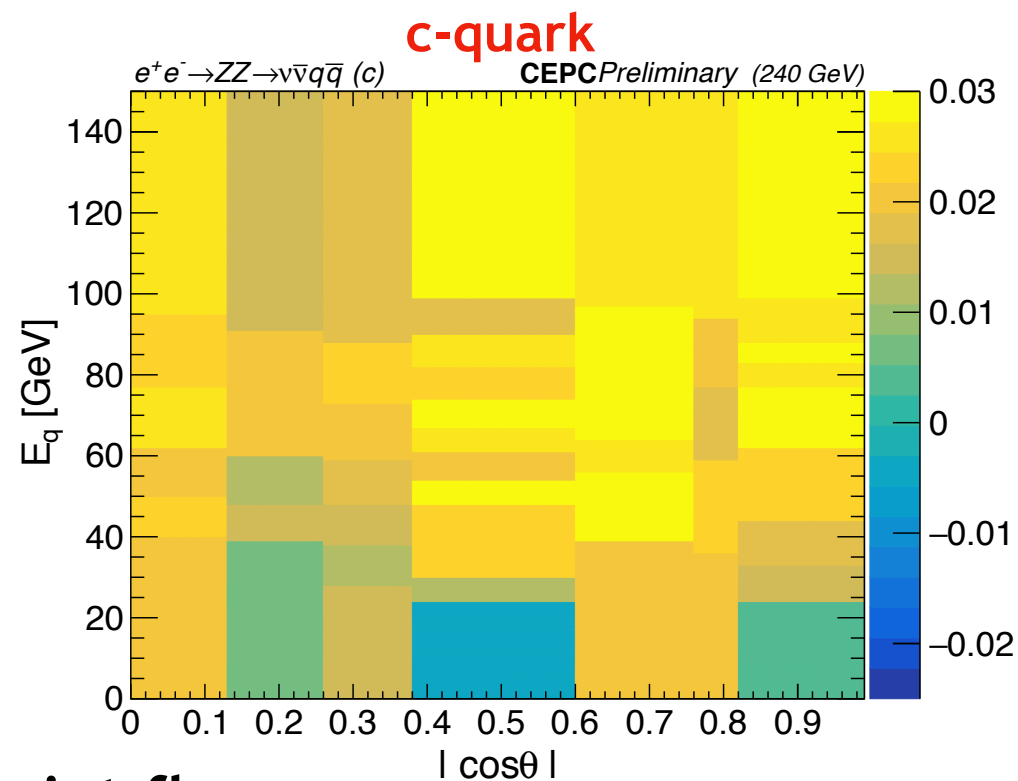
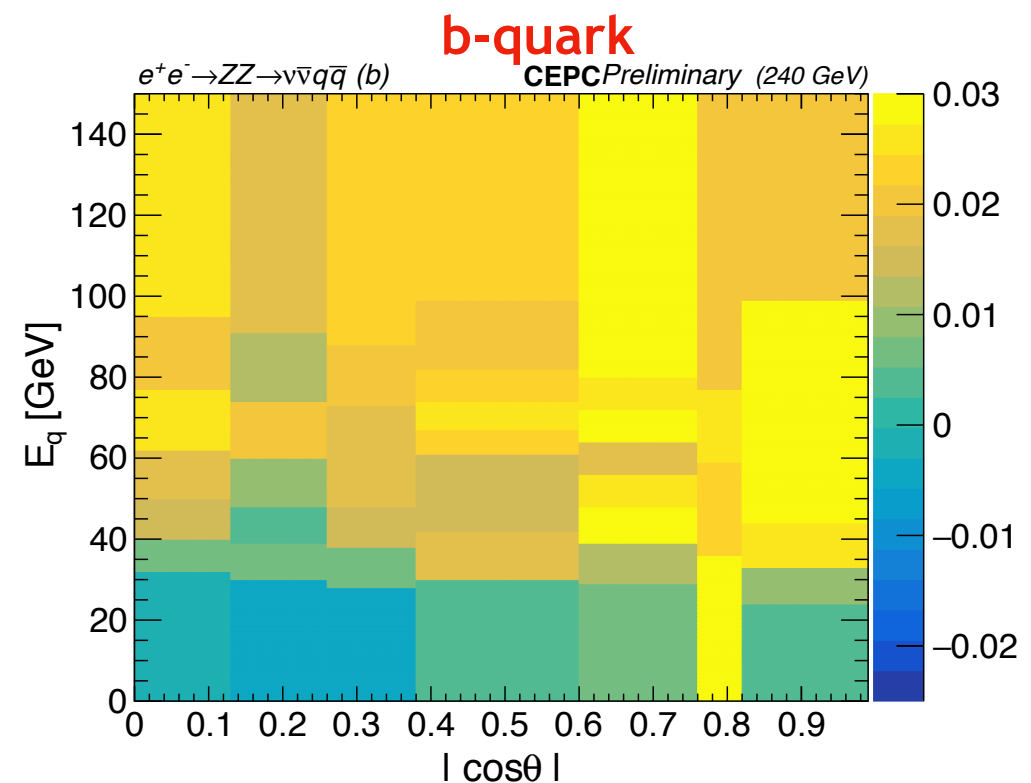
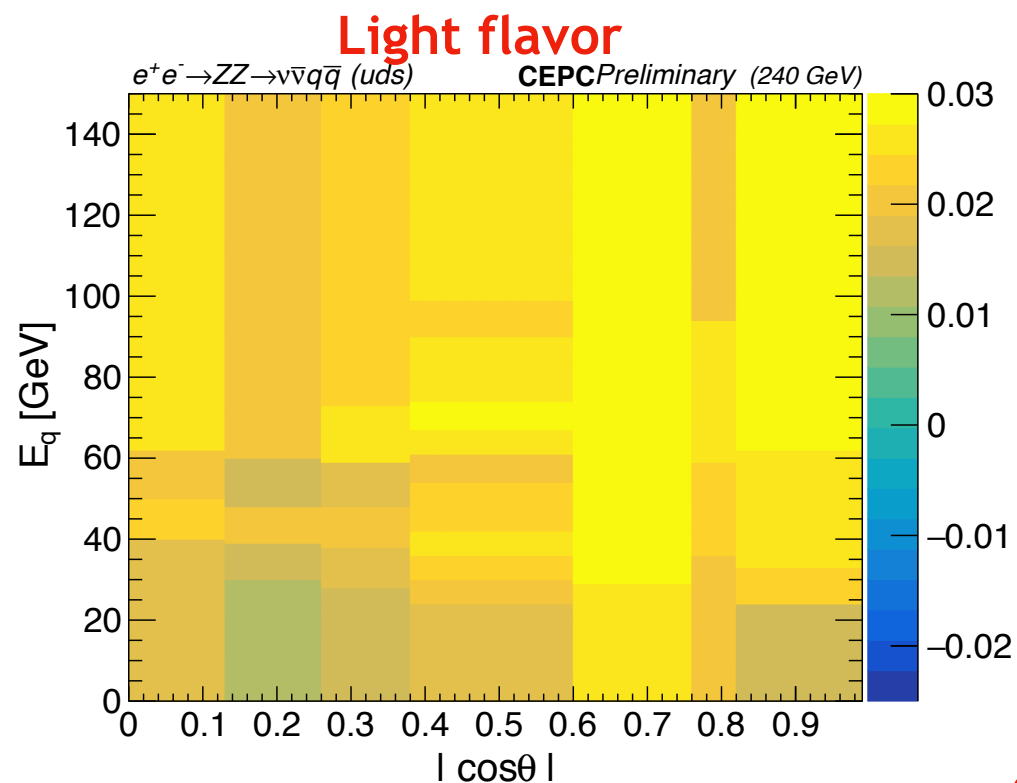
Reco-MCP



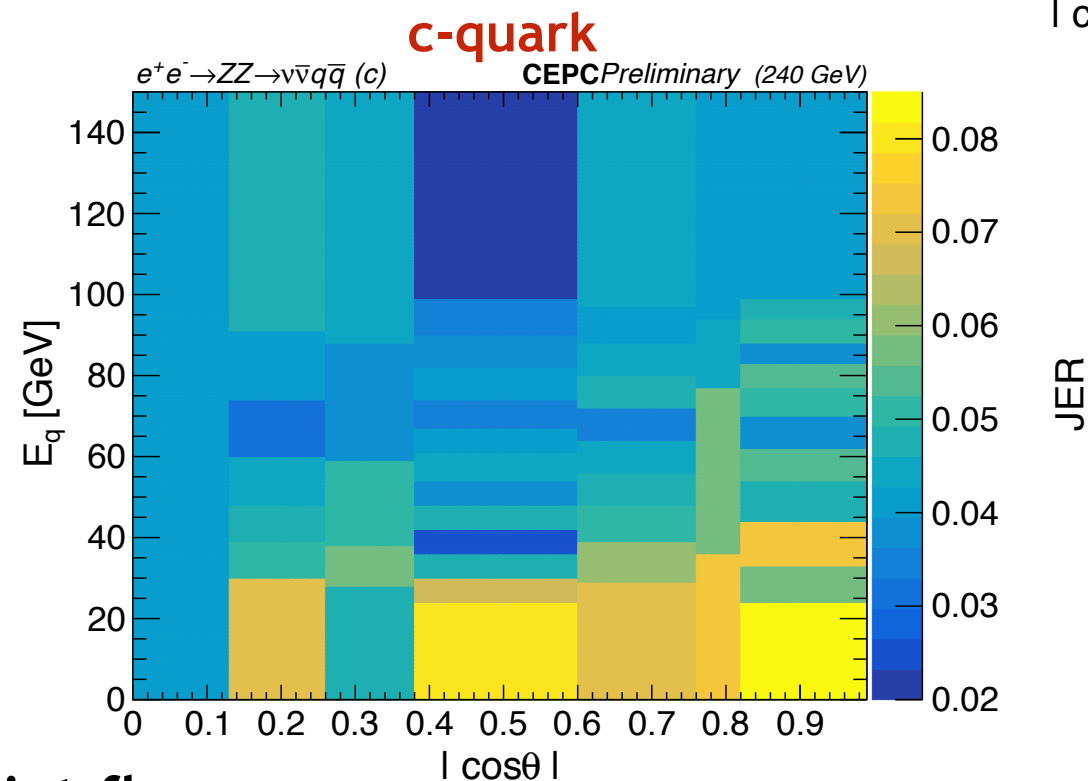
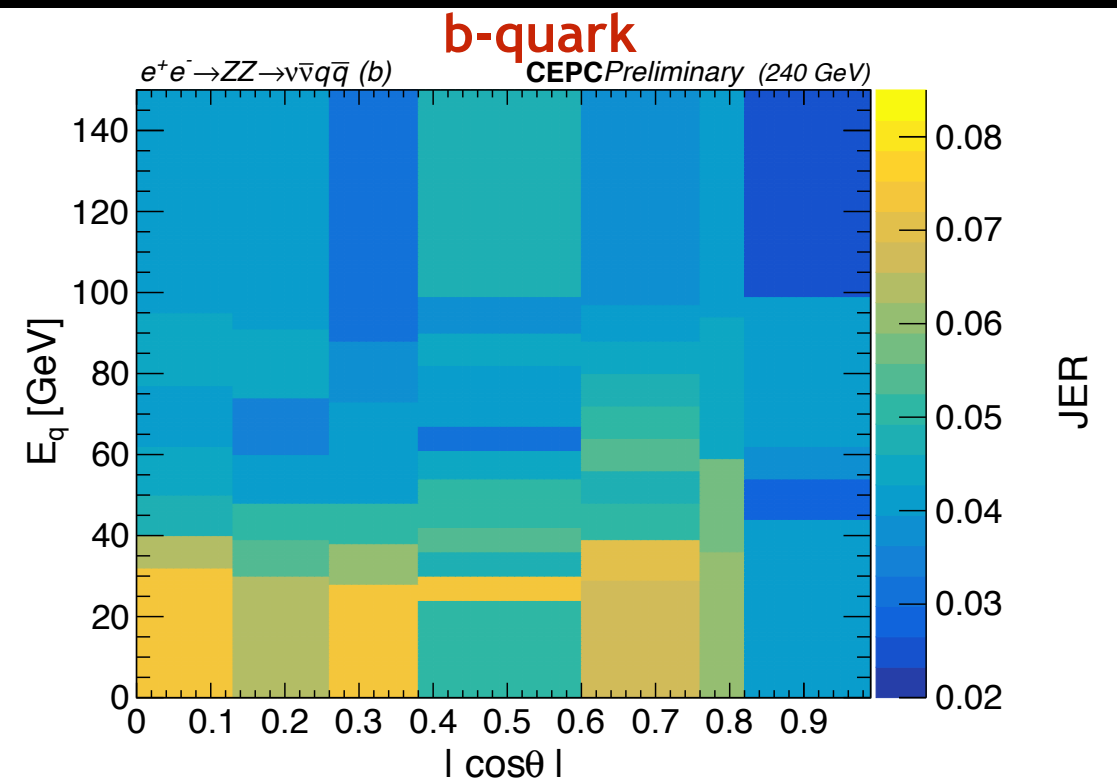
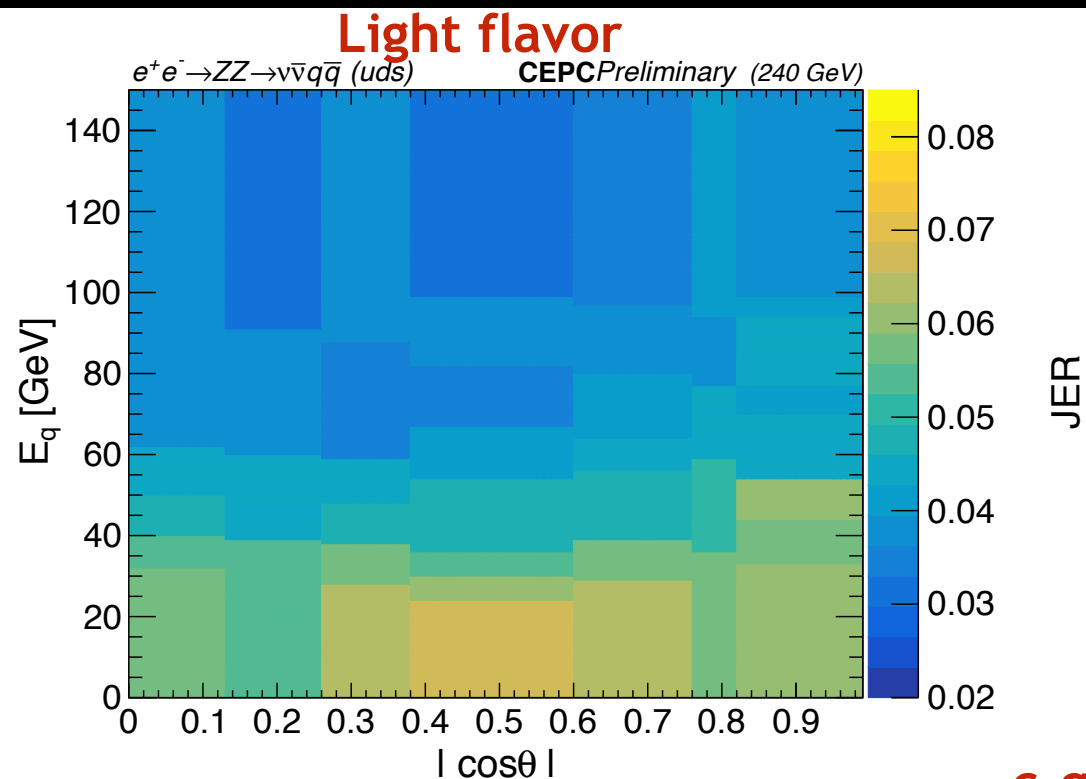
- JER/JES between Reco jet and MCP would combine the effects of two previous stages.



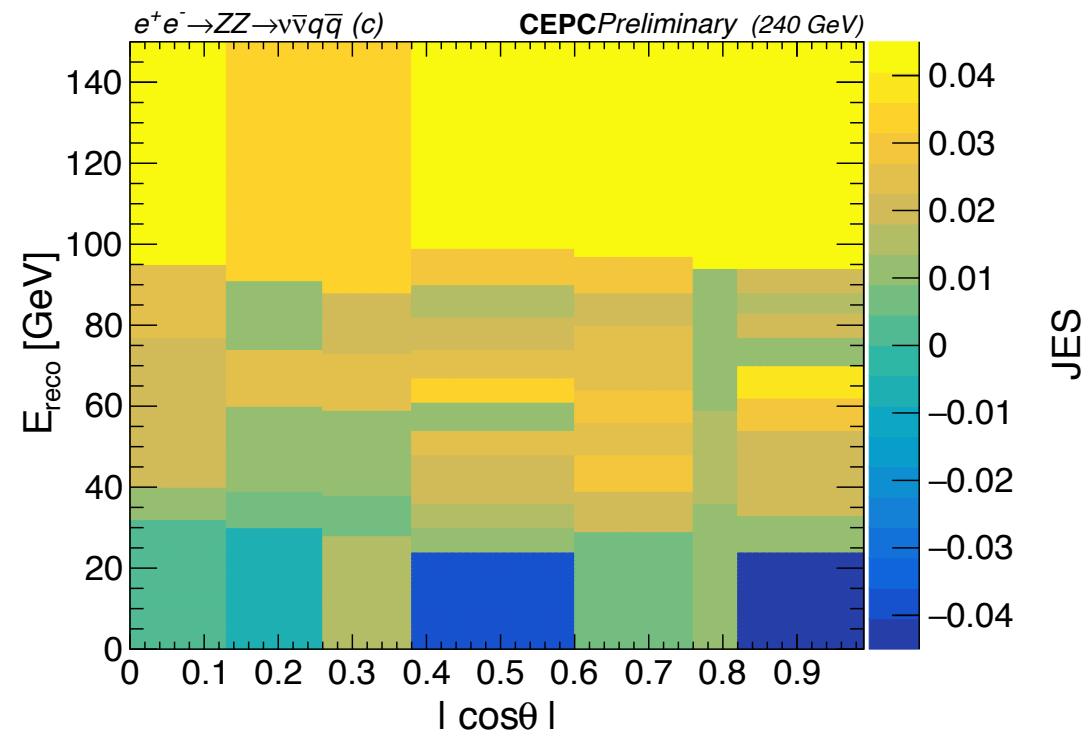
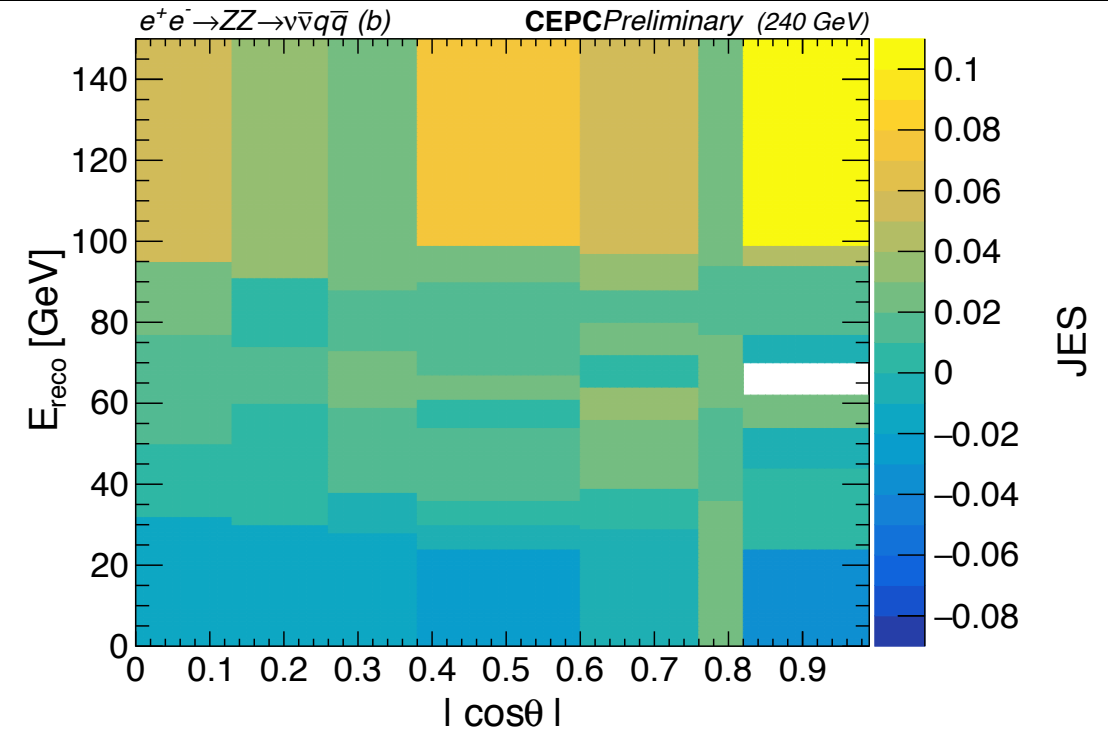
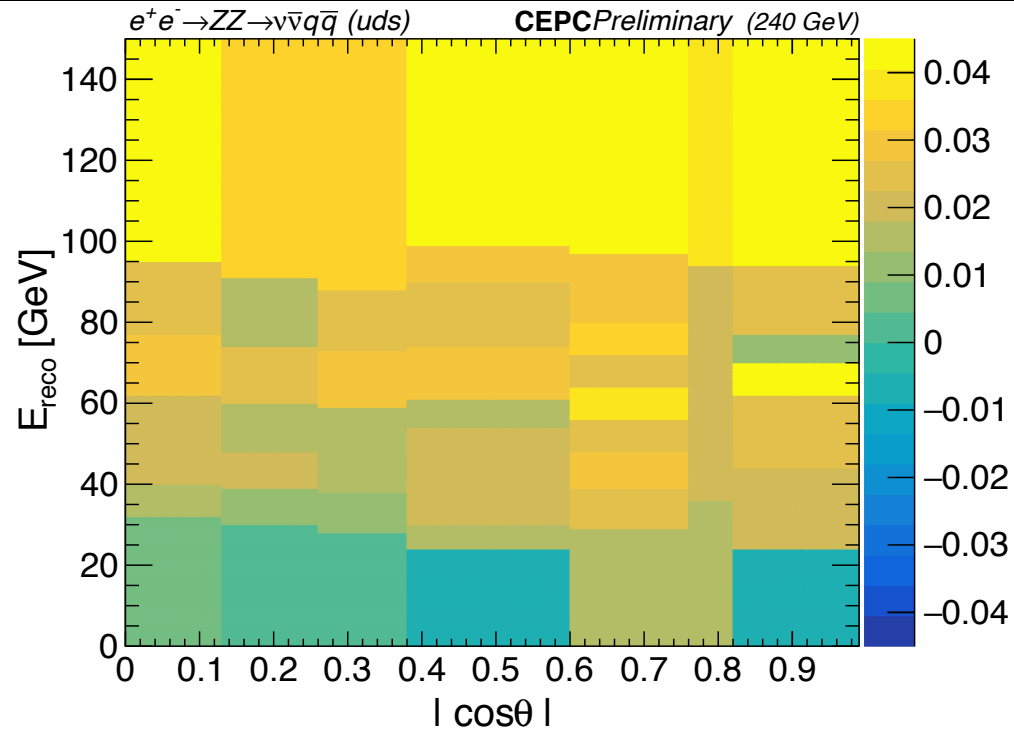
JES in Phase Space

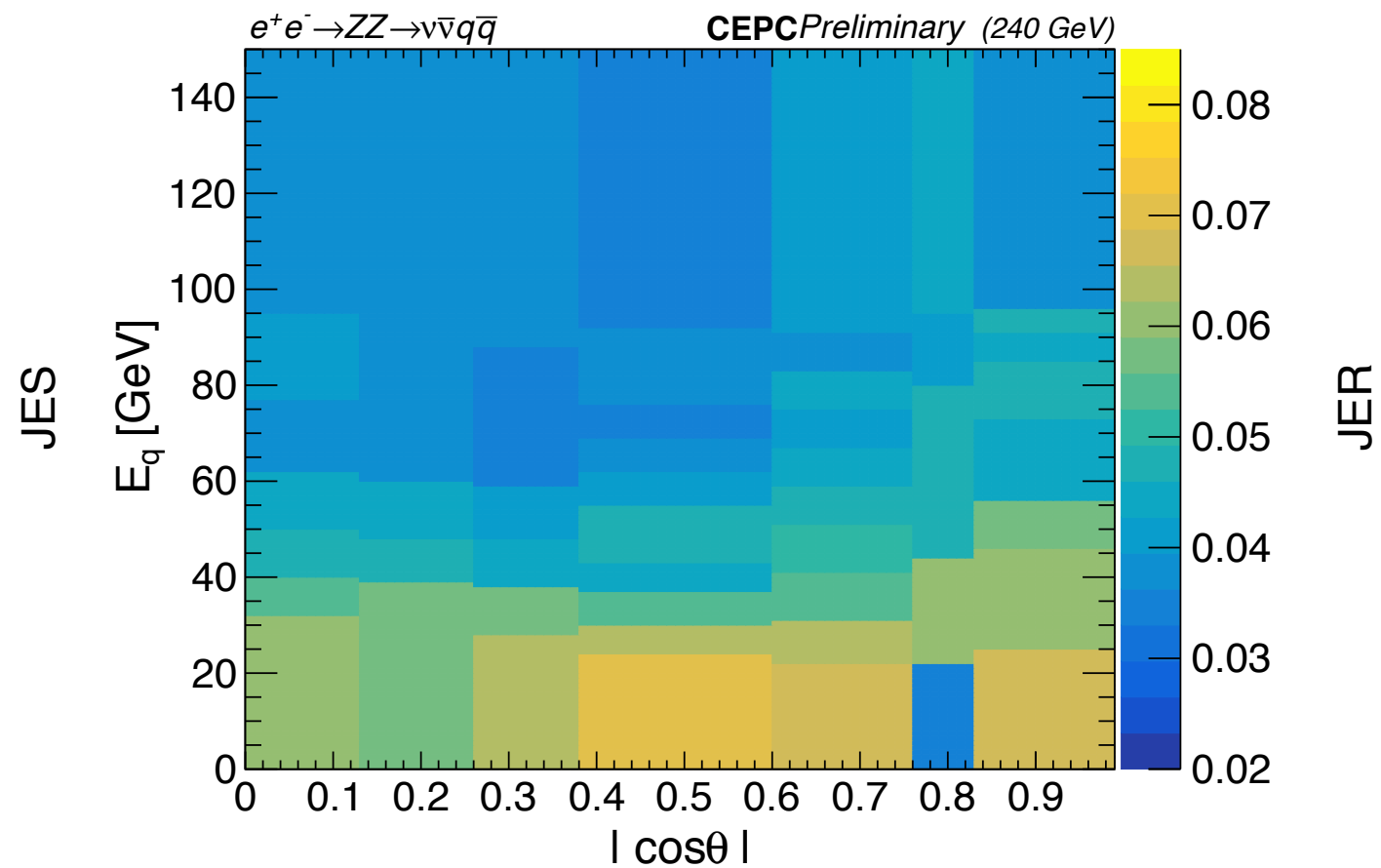
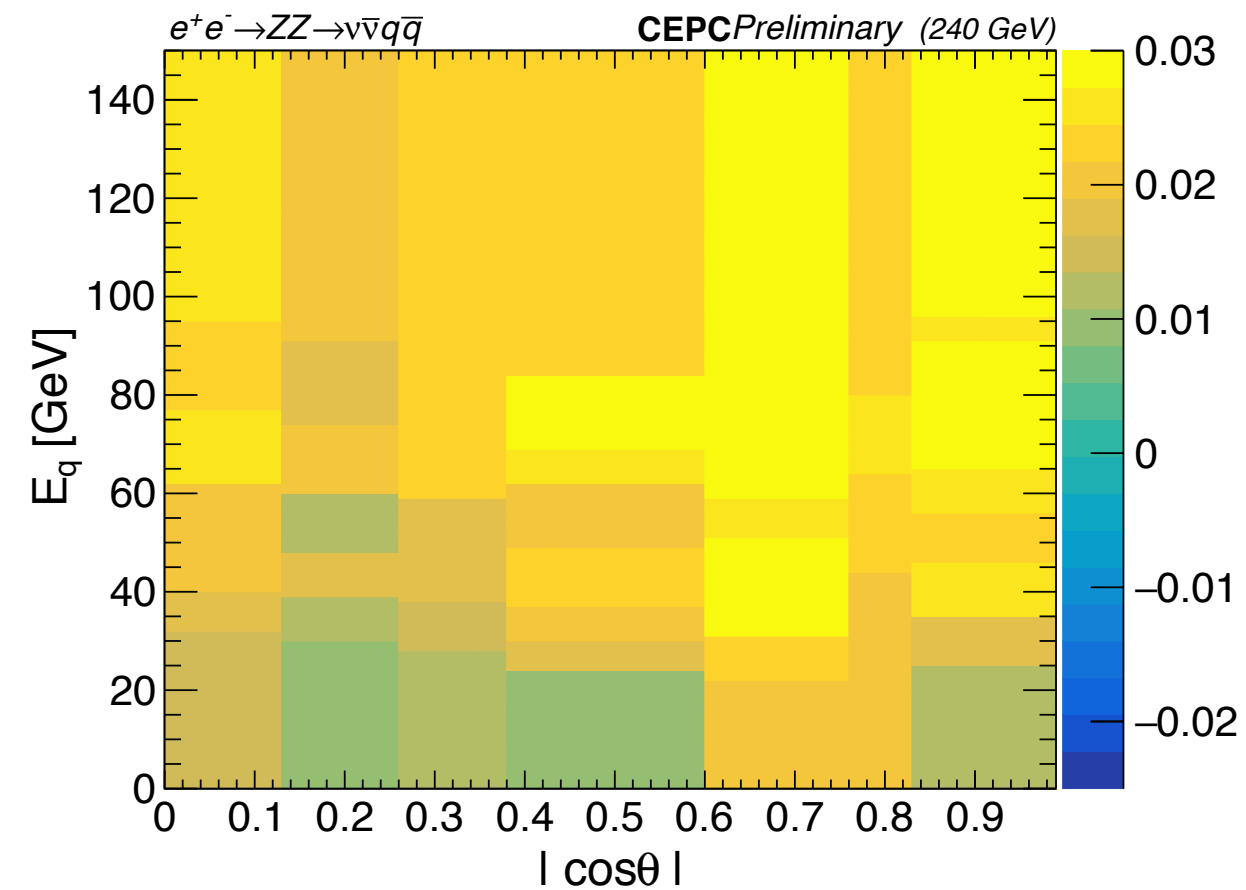


- JES also depends on the jet flavor.
- Light flavor jet has higher energy deviation.



- JER also depends on the jet flavor.
- Higher jet energy and within central region of barrel, JER has impressive performance.

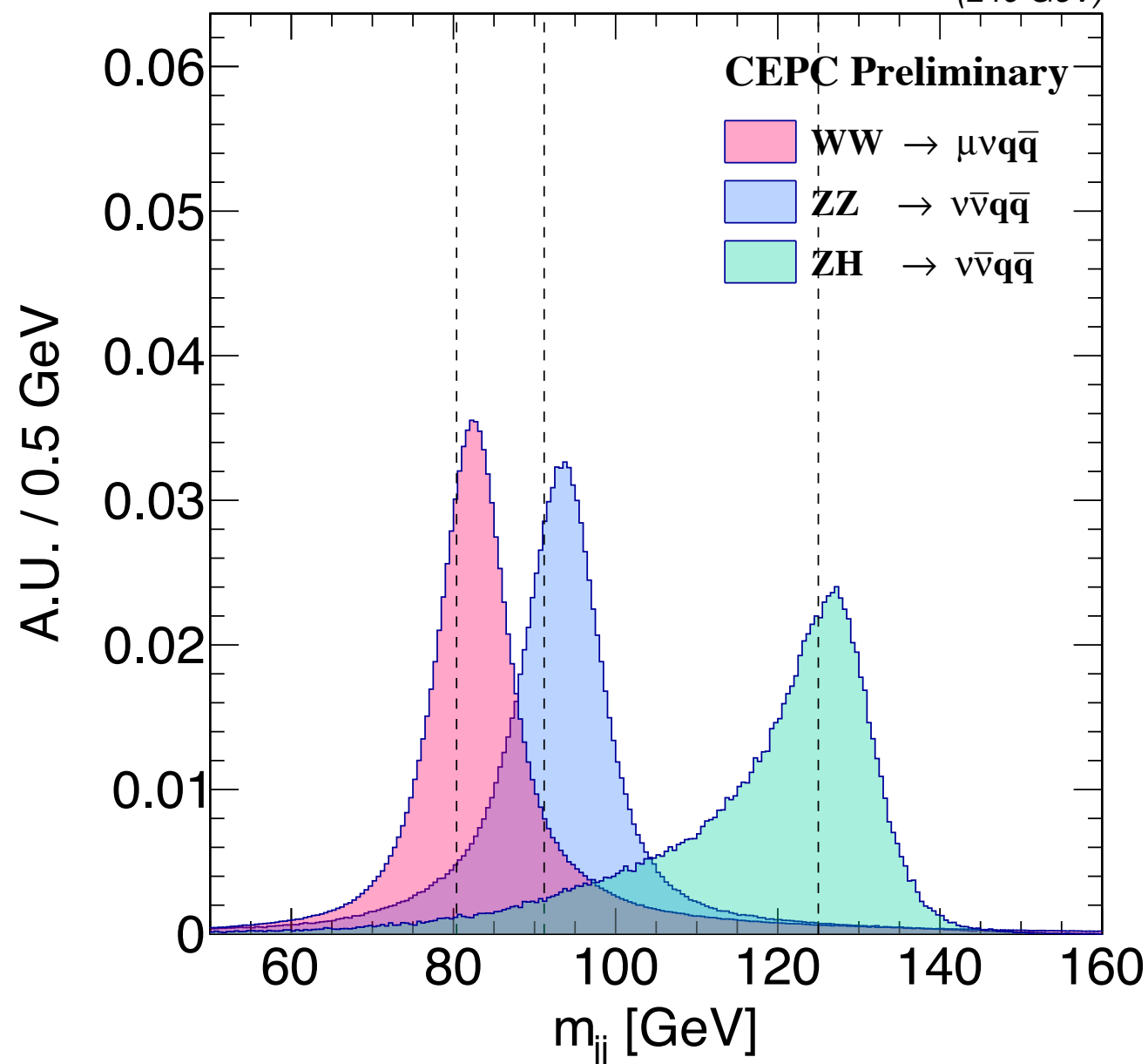




Jet Energy Calibration

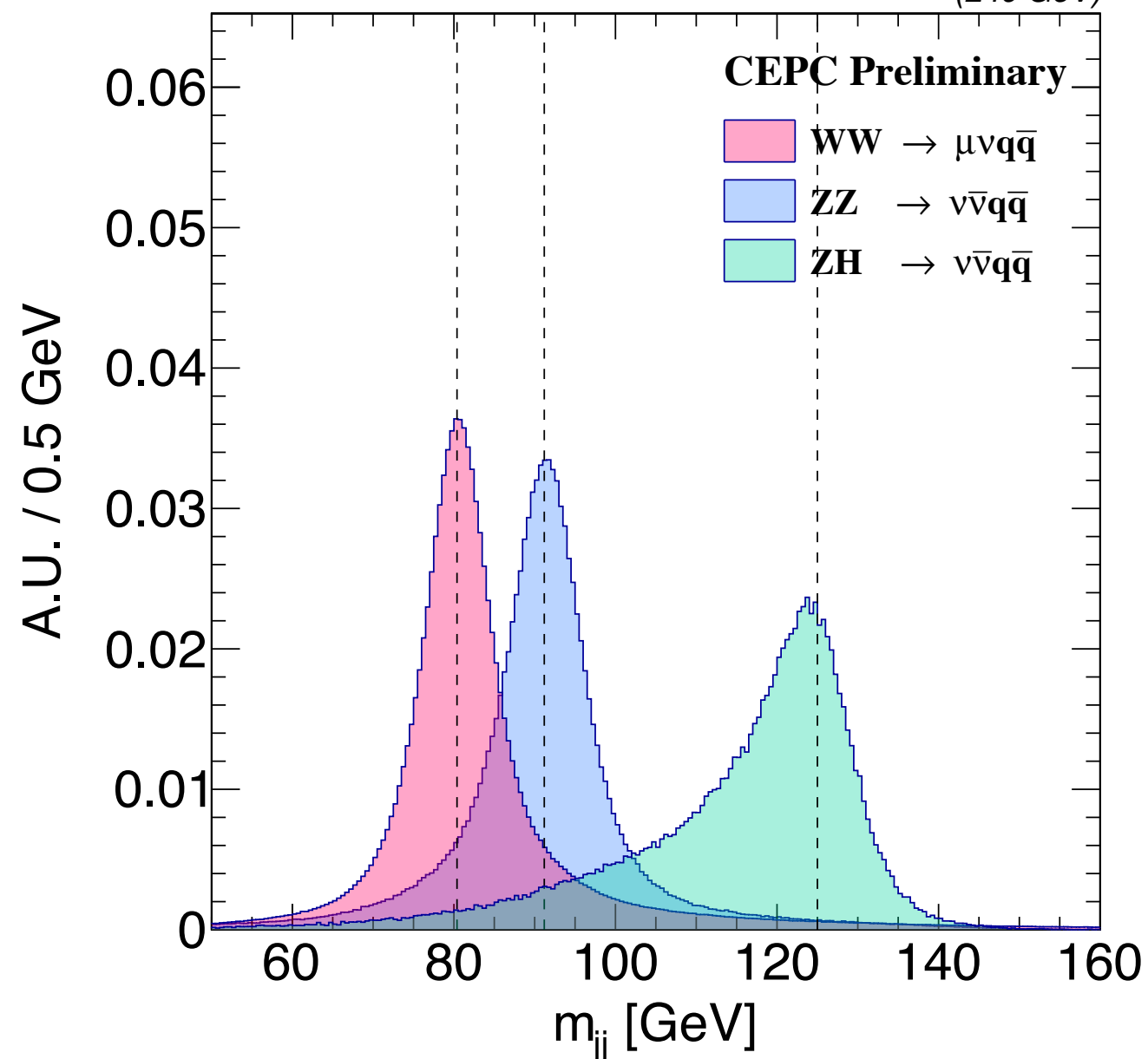
Before Calibration

(240 GeV)



After Calibration

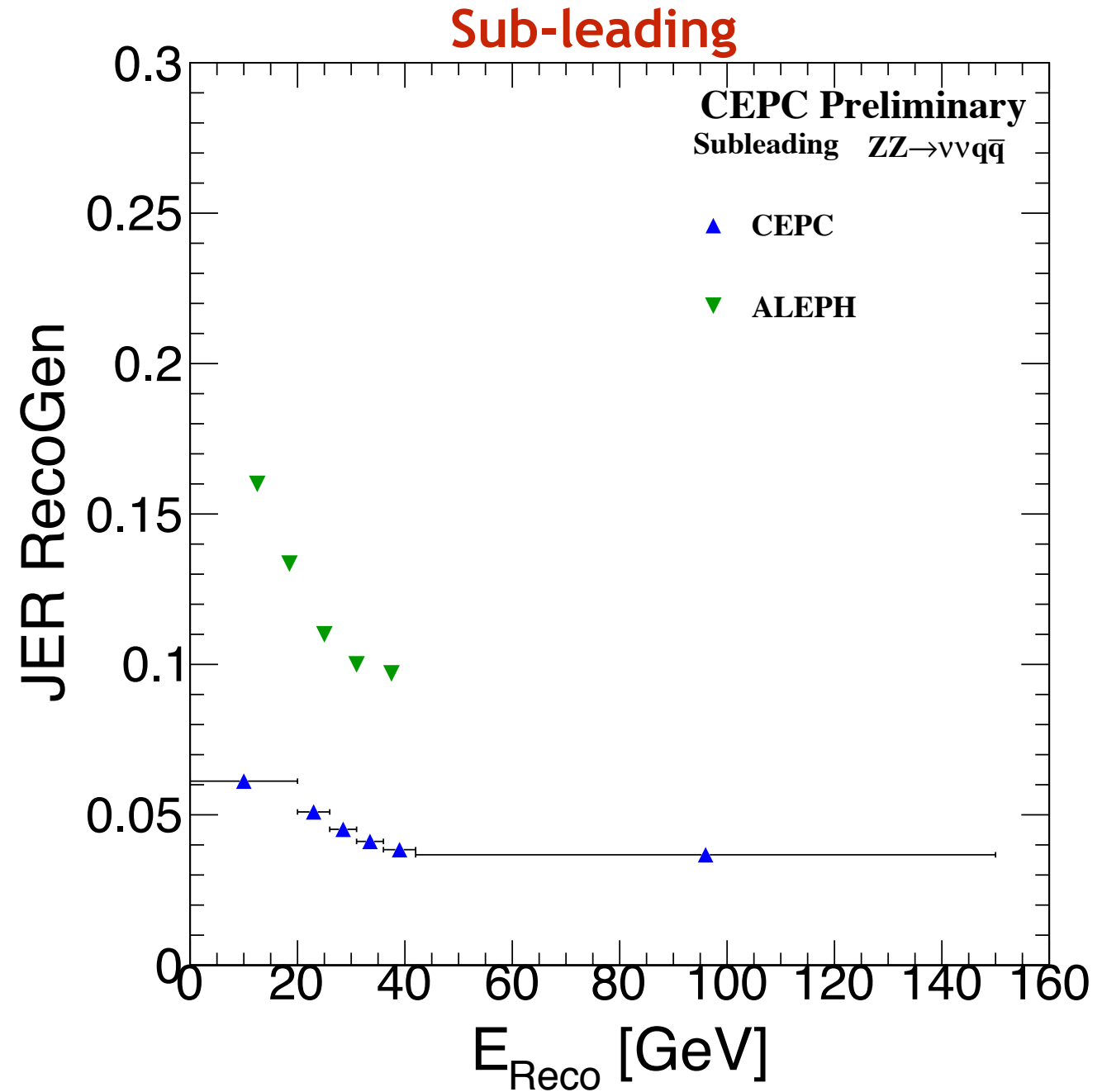
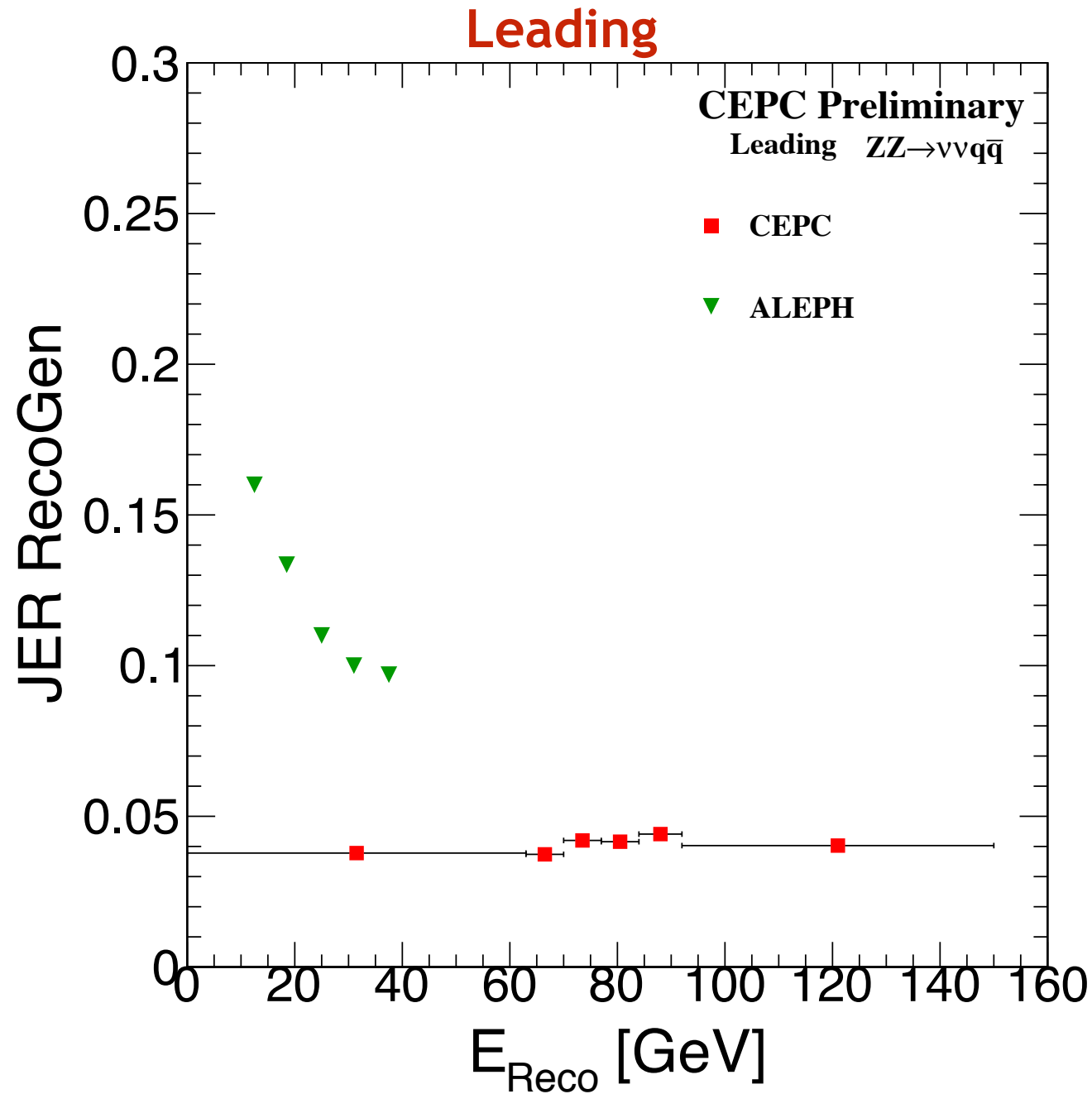
(240 GeV)



- Since the double-counting effect, jet energy would be overestimated.
- According to MC true energy and $\cos\theta$ distribution, JES can be used to calibrate the dijet invariant mass back to the value we put into simulation.
- After calibration, boson mass resolution is improved about **1%**.

m_W (GeV)	m_Z (GeV)	m_H (GeV)	Jets / PFOs	wi/wo Clean	wi/wo Cali
82.66 ± 3.54	93.69 ± 3.89	127.48 ± 4.93	Jets	0	0
82.79 ± 3.34	93.95 ± 3.48	127.31 ± 4.54	Jets	1	0
80.72 ± 3.46	91.67 ± 3.77	125.02 ± 5.11	Jets	0	1
80.82 ± 3.23	91.76 ± 3.39	124.39 ± 4.39	Jets	1	1
82.63 ± 3.53	93.69 ± 3.89	127.57 ± 4.80	PFOs	0	0
82.77 ± 3.32	93.90 ± 3.54	127.83 ± 4.50	PFOs	1	0

Compare with ALEPH at LEP



■ Our JER is better than ALEPH.