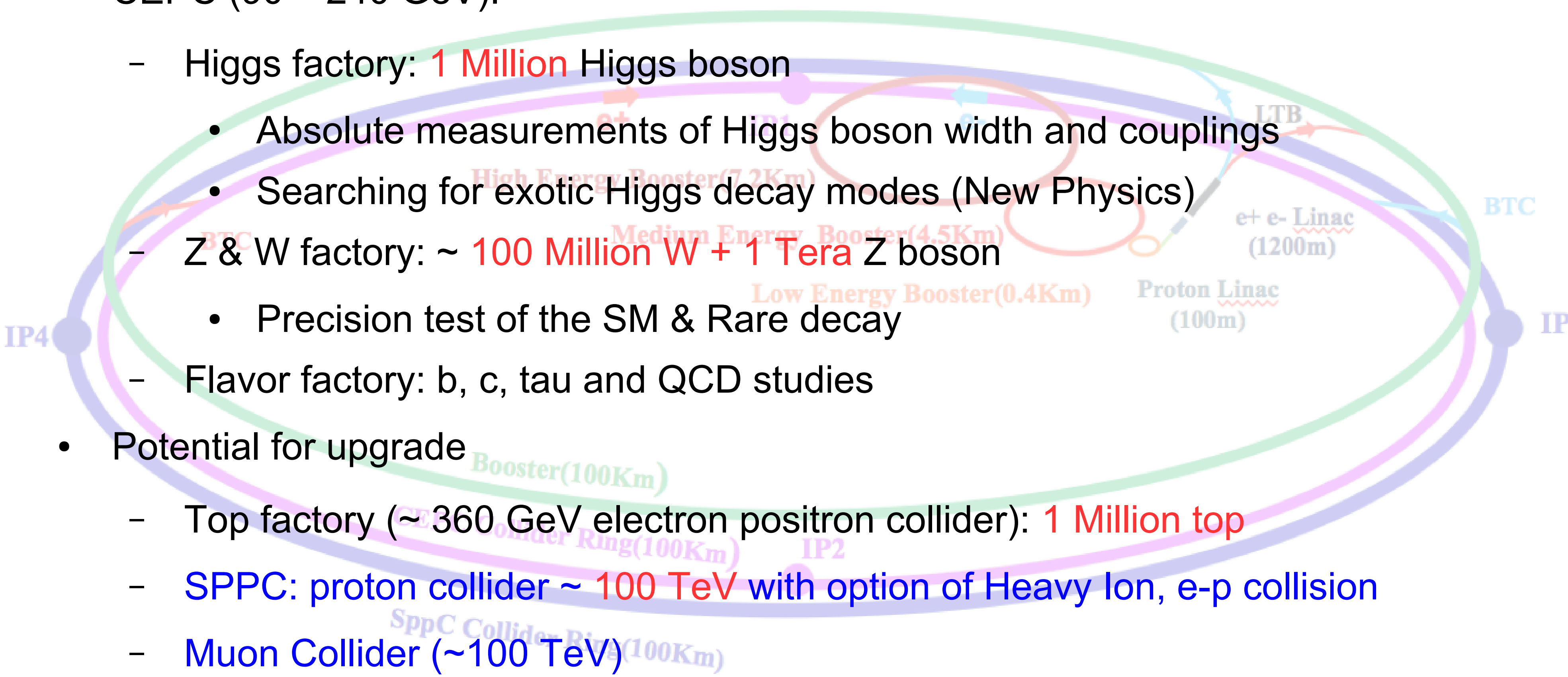
Jets Reconstruction at the CEPC

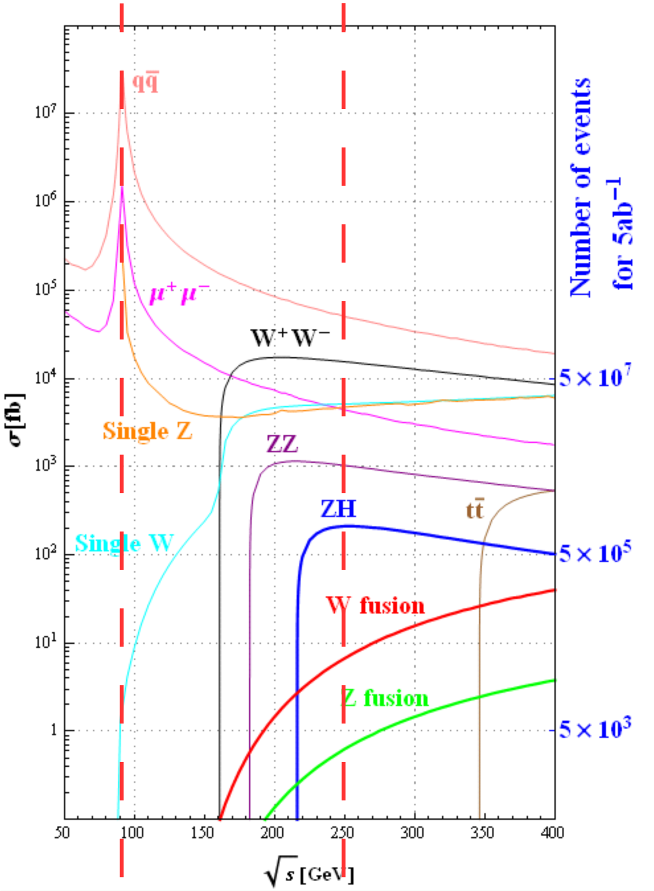
Manqi Ruan

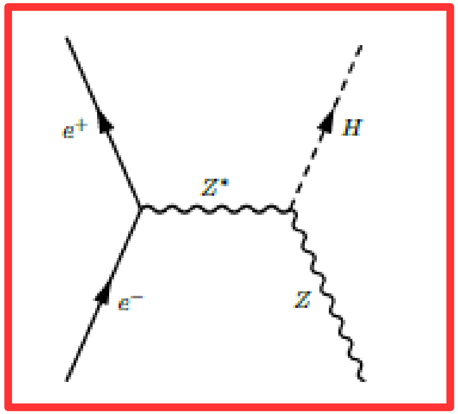
Key parameters of the CEPC-SPPC

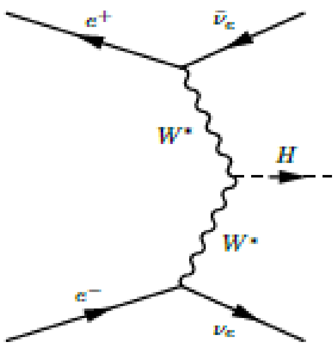
- Tunnel ~ 100 km
- CEPC (90 – 240 GeV):
  - Higgs factory: 1 Million Higgs boson
    - Absolute measurements of Higgs boson width and couplings
    - Searching for exotic Higgs decay modes (New Physics)
  - Z & W factory: ~ 100 Million W + 1 Tera Z boson
    - Precision test of the SM & Rare decay
  - Flavor factory: b, c, tau and QCD studies
- Potential for upgrade
  - Top factory (~ 360 GeV electron positron collider): 1 Million top
  - SPPC: proton collider ~ 100 TeV with option of Heavy Ion, e-p collision
  - Muon Collider (~100 TeV)
- A gigantic leap after the LHC: boost the precision on all frontiers (especially Higgs & EW) by order of magnitudes + enormous potential to discover New Physics Principle & reveal fore-mentioned puzzles

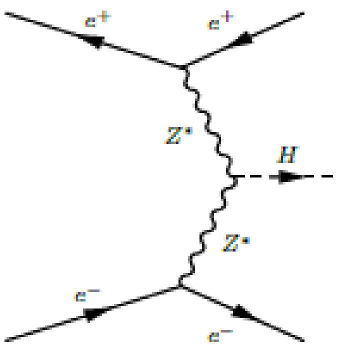


Higgs @ CEPC









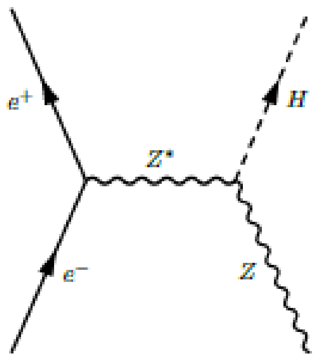
Process	Cross section	Events in 5 ab^{-1}
Higgs boson production, cross section in fb		
$e^+e^- \to ZH$	212	1.06×10^6
$e^+e^- \to \nu\bar{\nu}H$	6.72	3.36×10^4
$e^+e^- \to e^+e^-H$	0.63	3.15×10^3
Total	219	1.10×10^6

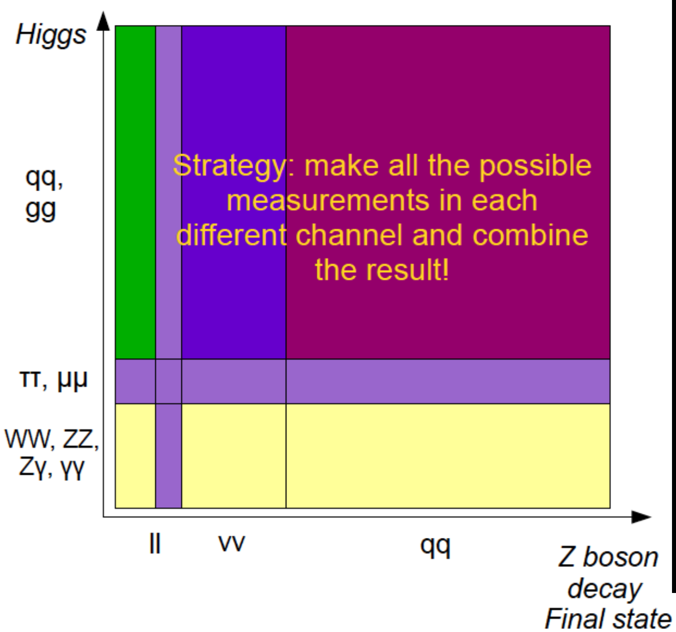
S/B ~ 1:100 - 1000

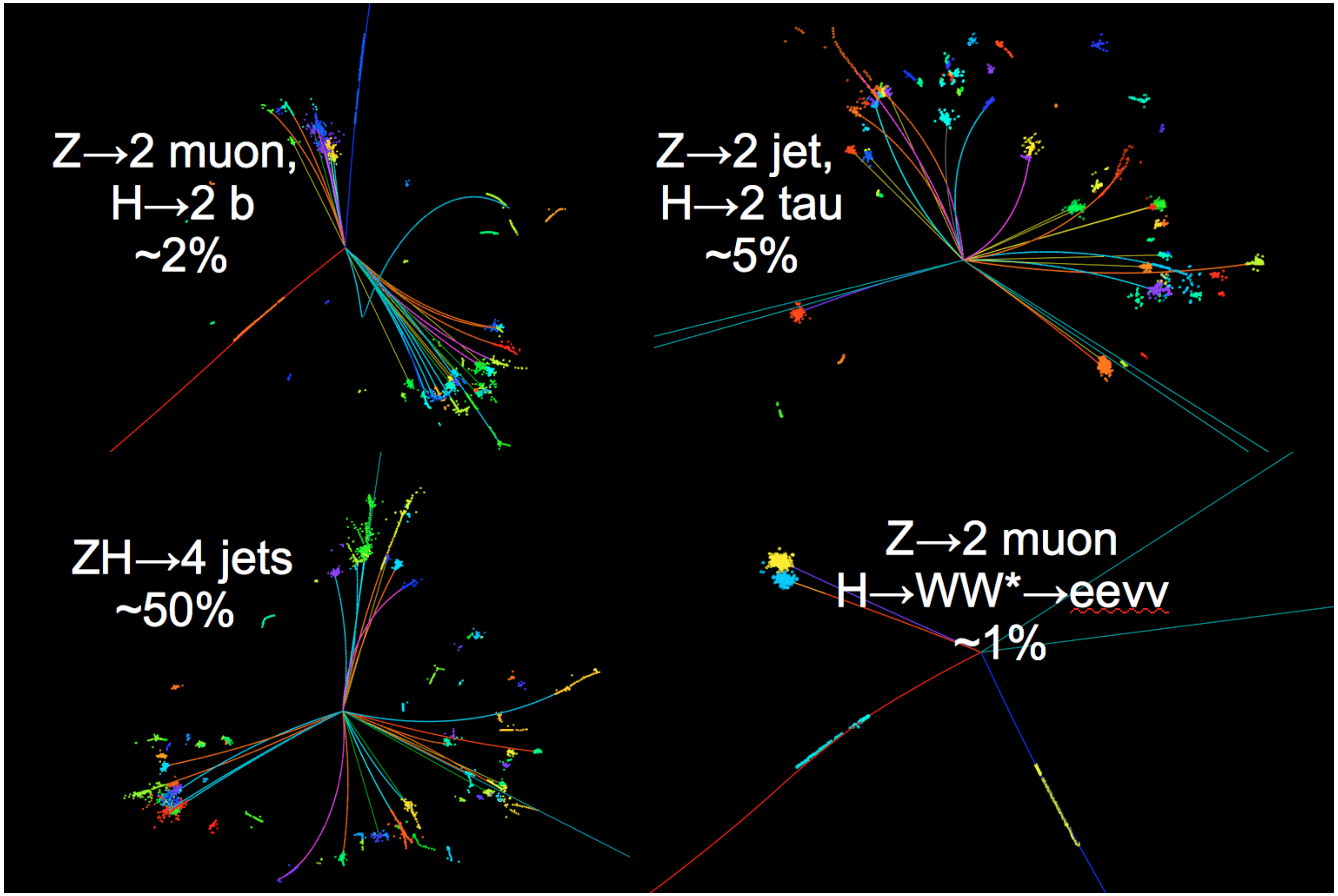
Observables: Higgs mass, CP, σ(ZH), event rates (σ(ZH, vvH)*Br(H→X)), Diff. distributions

Derive: Absolute Higgs width, branching ratios, couplings

Physics Requirements





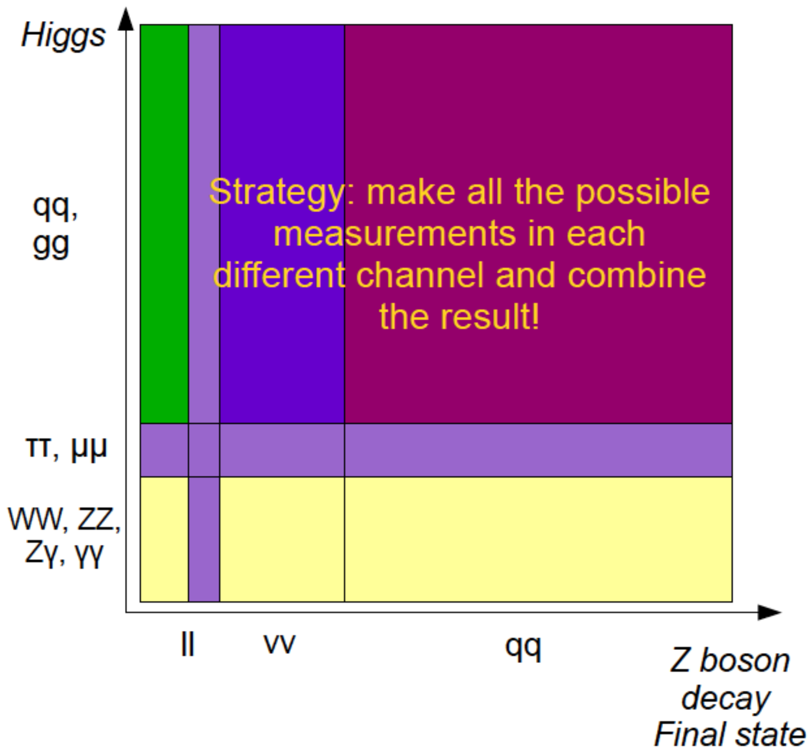


Detector:

To reconstruct all the physics objects with high efficiency, purity & resolution
Homogenous & Stable enough to control the systematic

Jets at Higgs Signal

- SM Higgs
  - **0 jets: 3%**
    - Z→ll, νν (30%); H→0 jets (~10%, ττ, μμ, γγ, γZ/WW/ZZ→leptonic)
  - **2 jets (+n with gluon emission...): 32%**
    - Z→qq, H→0 jets. 70%*10% = 7%
    - Z→ll, νν; H→2 jets. 30%*70% = 21%
    - Z→ll, νν; H→WW/ZZ→semi-leptonic. 3.6%
  - **4 jets: 55%**
    - Z→qq, H→2 jets. 70%*70% = 49%
    - Z→ll, νν; H→WW/ZZ→4 jets. 30%*15% = 4.5%
  - **6 jets: 11%**
    - Z→qq, H→WW/ZZ→4 jets. 70%*15% = 11%



- 97% of the SM Higgsstrahlung Signal involves Jets
- 66% need color-singlet identification: grouping the hadronic final sate particles into color-singlets (Z, H, W, gamma, ...).

Jets at Higgs Signal

- SM Higgs
  - **0 jets: 3%**
    - Z→ll, νν (30%); H→0 jets
  - **2 jets: 32%**
    - Z→2 jets, H→0 jets. 7%
    - Z→0 jets; H→2 jets. 21%
    - Z→0 jets; H→VV→semi-leptonic. 3.6%

- 1/3 of the Higgsstrahlung events
  - Have access to all SM Higgs decay modes
  - Doesn't need color singlet identification: No hadronic final states, or only 1 color singlet thus naturally identified

---

  - **4 jets: 55%**
    - Z→2 jets, H→2 jets + X. 49%
    - Z→0 jets; H→WW/ZZ→4 jets. 4.5%
  - **6 jets: 11%**
    - Z→qq, H→WW/ZZ→4 jets. 11%

- 2/3 of the Higgsstrahlung events
  - Dominate statistic of H→bb, cc, gg, WW, ZZ, Zγ
  - Color singlet identification – **potentially a leading systematic, huge impact**

- Jet clustering is essential for any measurements concerning jet direction (differential measurements)

Jets at other SM Processes

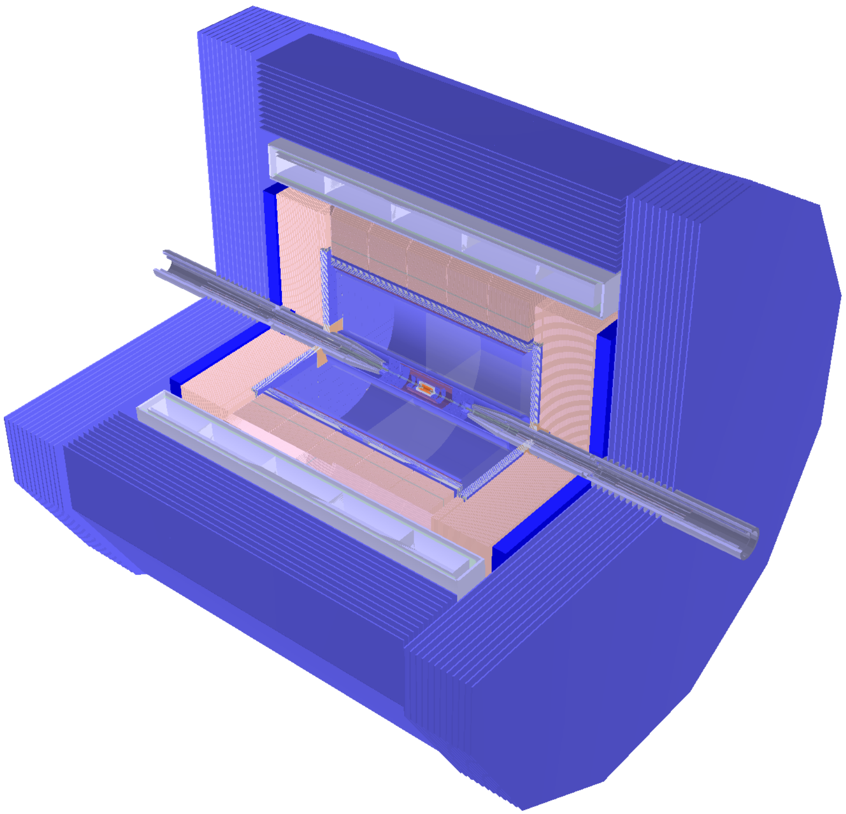
- 0 jets:
  - Di-photon events;
  - bhabha, ττ, μμ;
  - WWZ, leptonic;
- 2 jets: (+n with gluon emission...)
  - ee→qq(γ) (ISR return & full energy)
  - WW→semi-leptonic
  - Single W/Z events
  - Top pairs, WWZ, VBS (semi)-leptonic

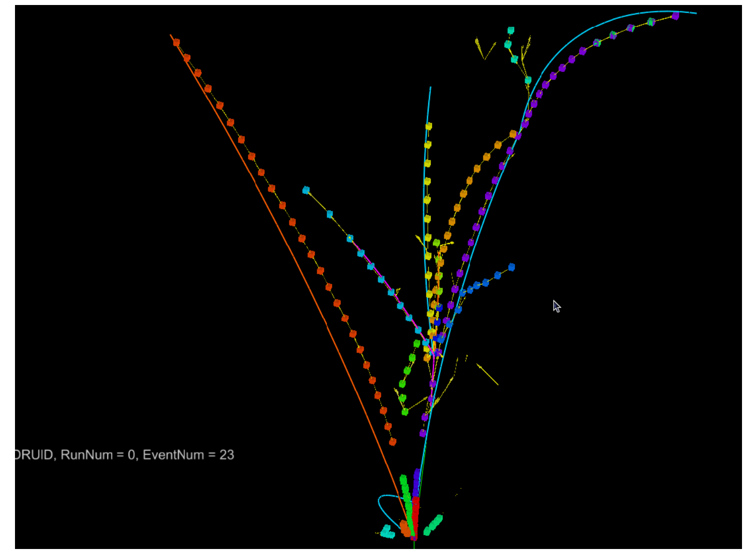
- The majority of the SM statistics doesn't need color singlet identification (0-1 color singlet goes into hadronic final states)
- Processes can be distinguished clearly
  - Benchmark eff*purity > 90%
- How can we made use of them?
  - W mass, TGC...
  - **And More?**

- 4 jets:
  - WW→Full hadronic
  - ZZ→Full hadronic
  - Top pairs, WWZ, semi-leptonic
  - VBS hadronic
- 6 jets: Top pairs, WWZ: full-hadronic

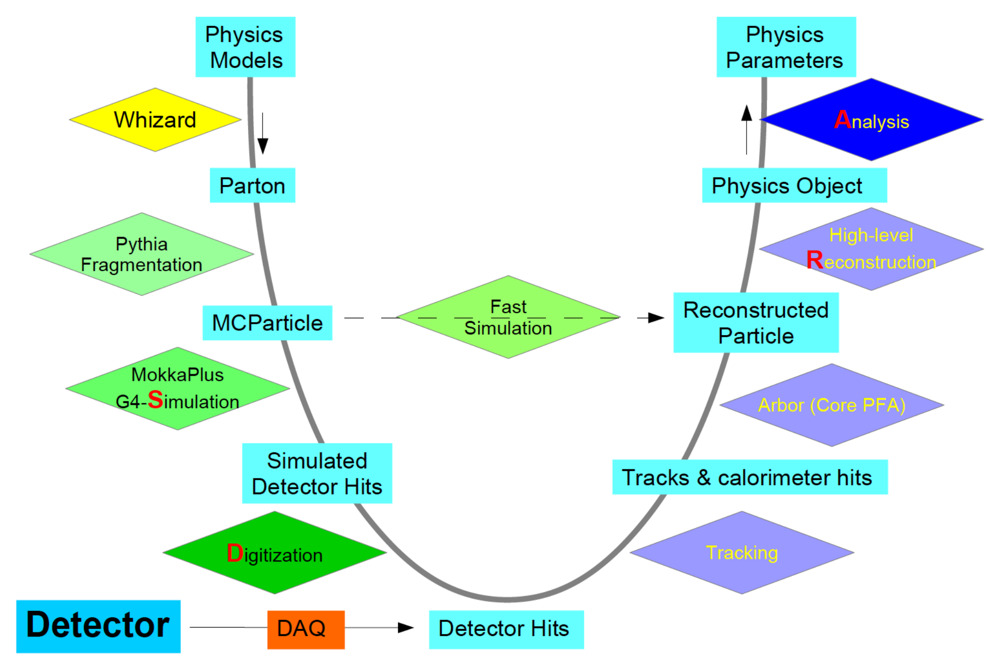
- At increasing sqrt(s), multi-jet events become significant
- Color singlet identification – **potentially a leading systematic, huge impact**
- *Demonstrated with full hadronic WW/ZZ separation at 240 GeV*
- *W boson identification at Top pair/WWZ events...*

CEPC Baseline: Particle Flow Oriented

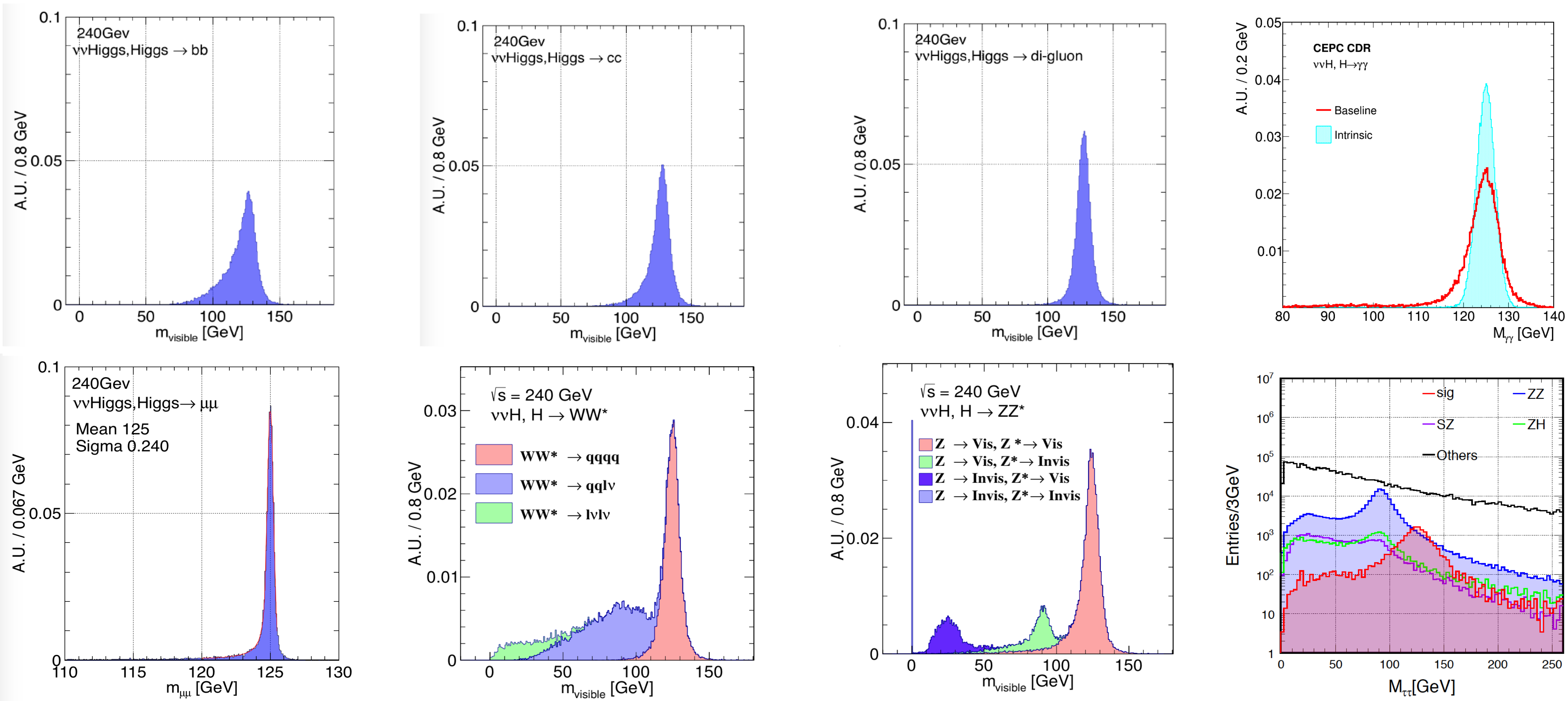




CEPC-SIMU-2017-001,
CEPC-SIMU-2017-002,
(DocDB id-167, 168, 173)



Higgs Signature at Baseline



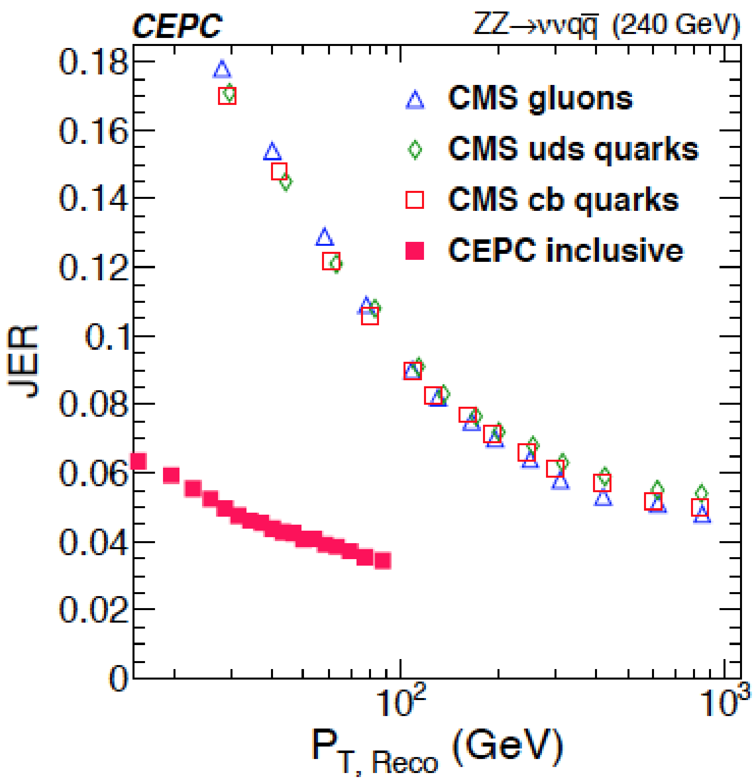
Clear Higgs Signature in all SM decay modes

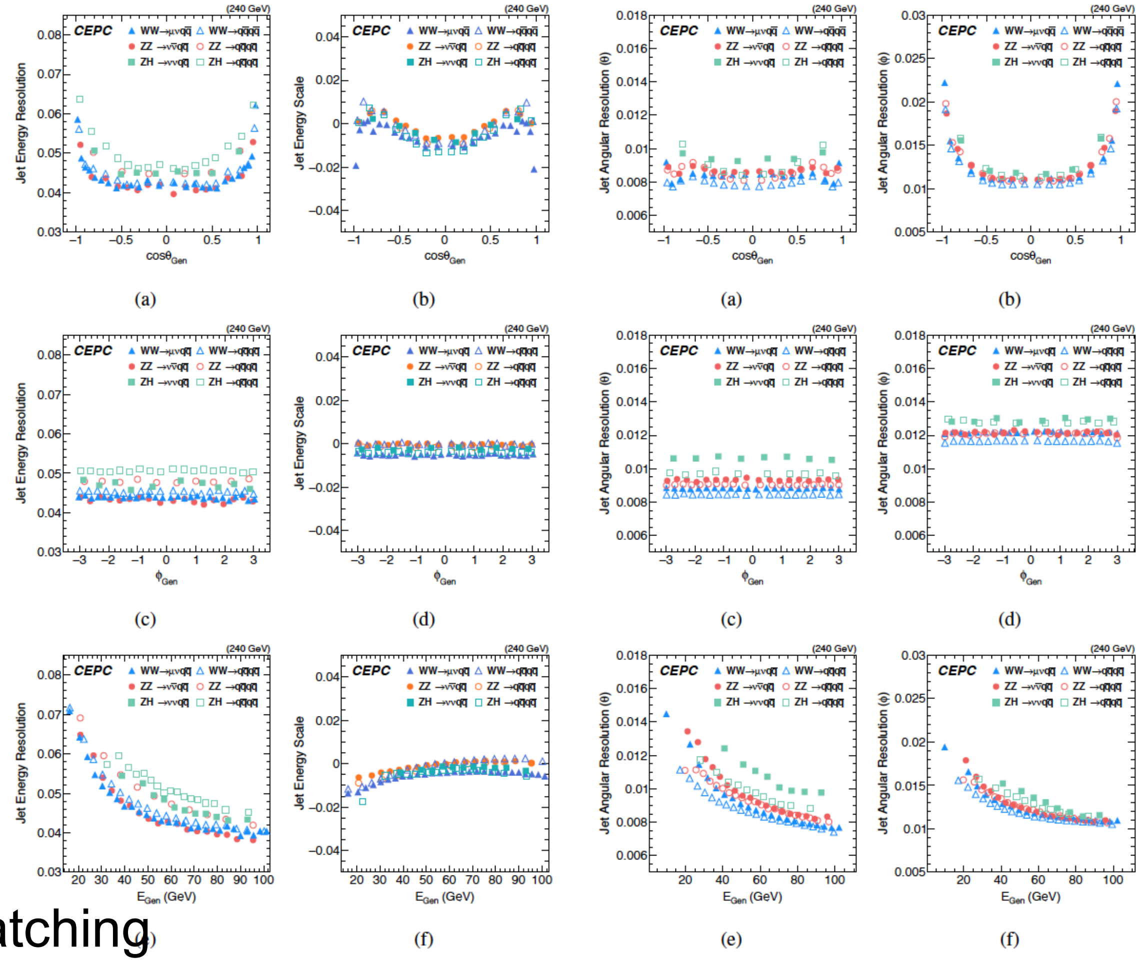
Massive production of the SM background (2 fermion and 4 fermions) at the full Simulation level

Right corner: di-tau mass distribution at qqH events using collinear approximation

Jet Energy Resolution & Scale

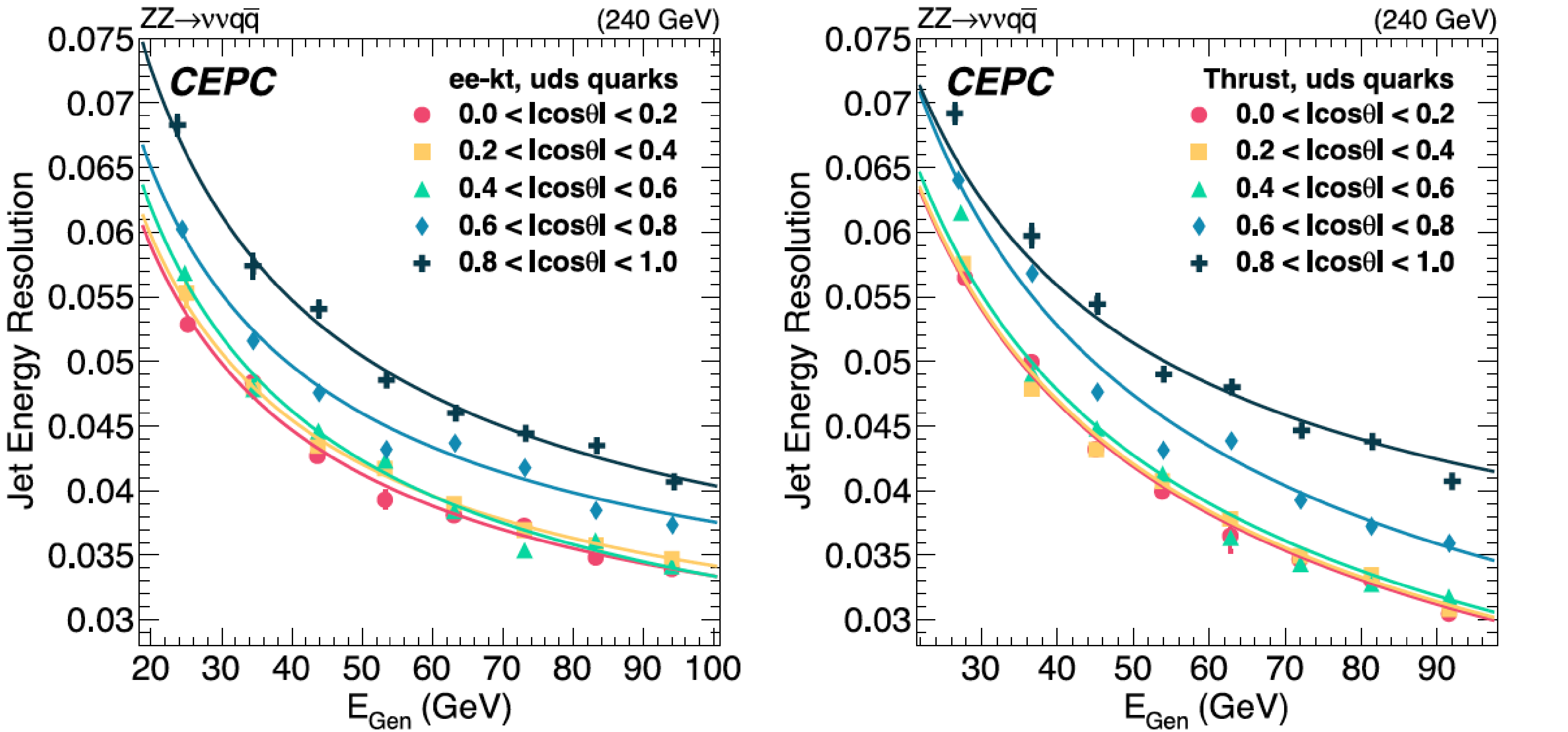
2021 JINST 16 P07037





- Conventional jet clustering + matching
  - JER: 2-4 times better than CMS @ 0 PU
  - JES: 0.1%/1% level with/without polar angle dependent calibration

Jet response: dependence on Jet clustering & matching...

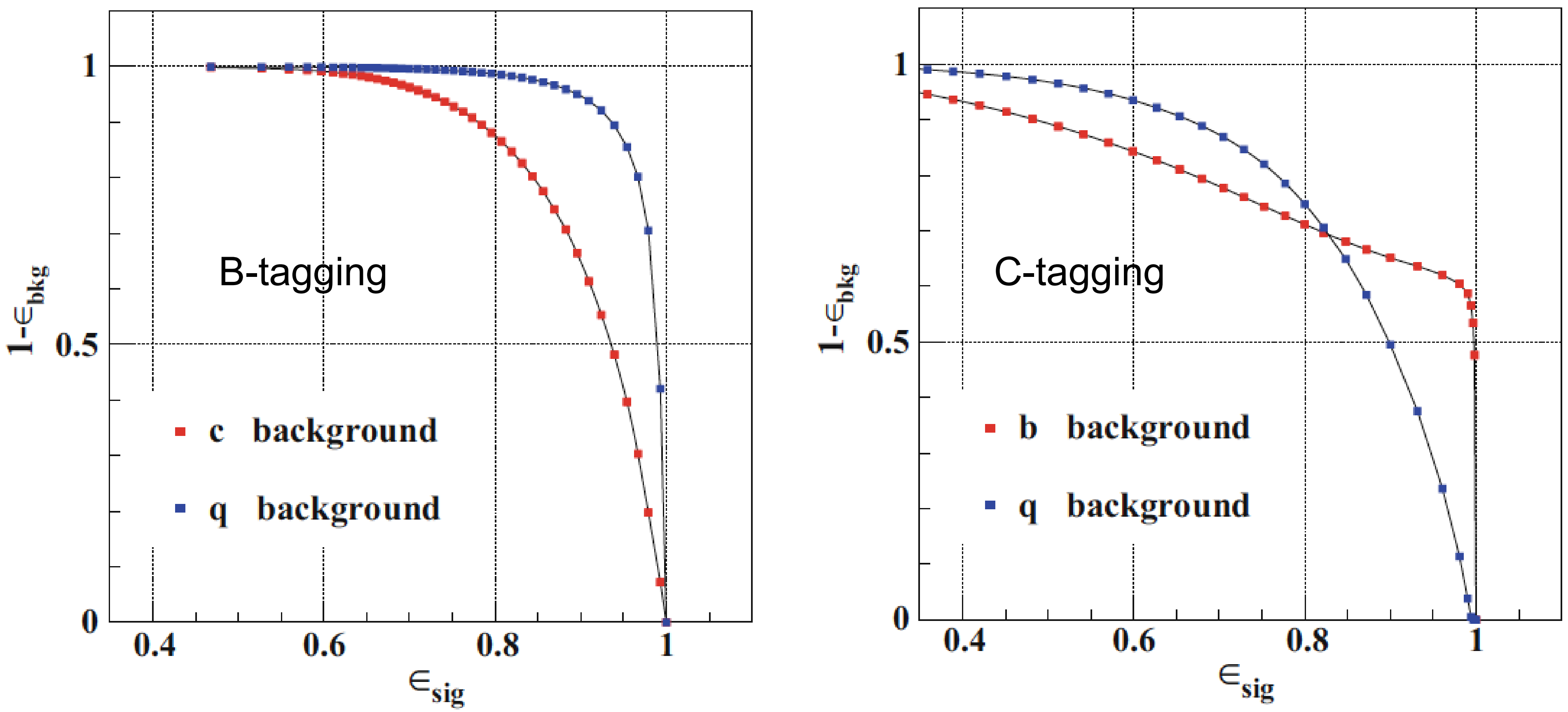


2021 JINST 16 P07037

(a) (b)

- Relative difference of ~10%

Jet Flavor Tagging



CEPC CDR Vol. 2

- B-tagging: eff*purity @ Z→qq: 70%
- C-tagging: eff*purity @ Z→qq: 40%

Latest update: See Libo Liao's talk

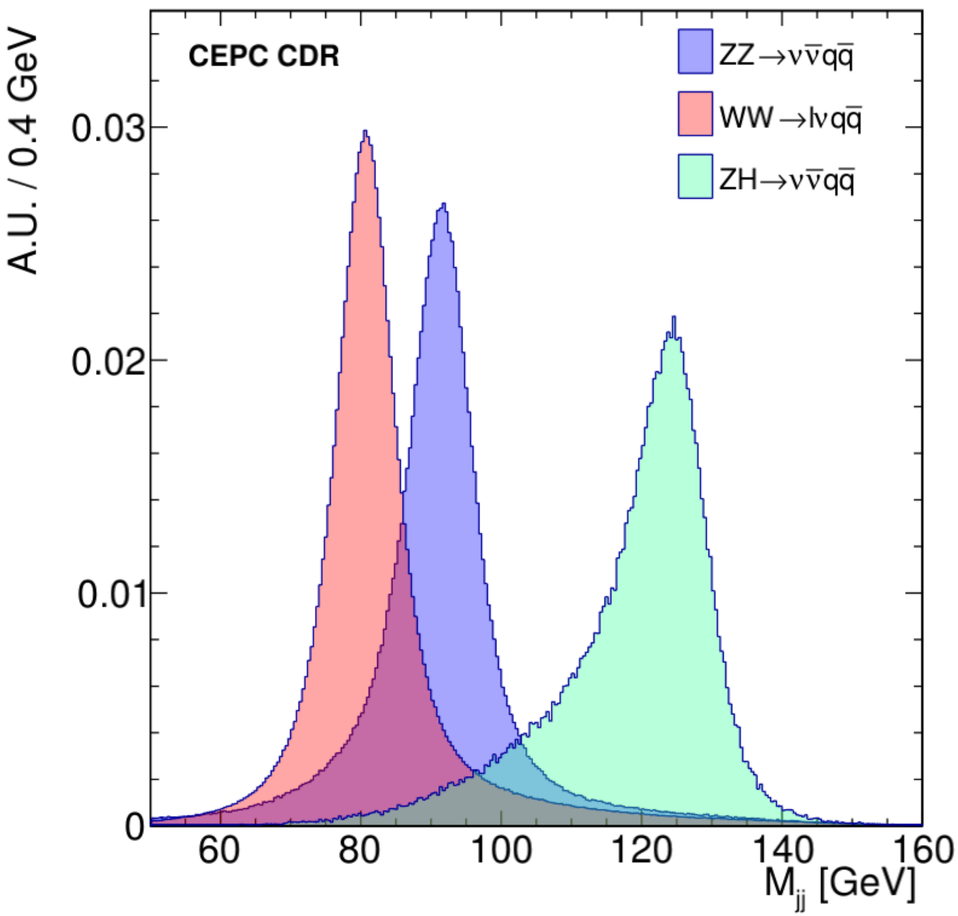
Key performances & Benchmarks

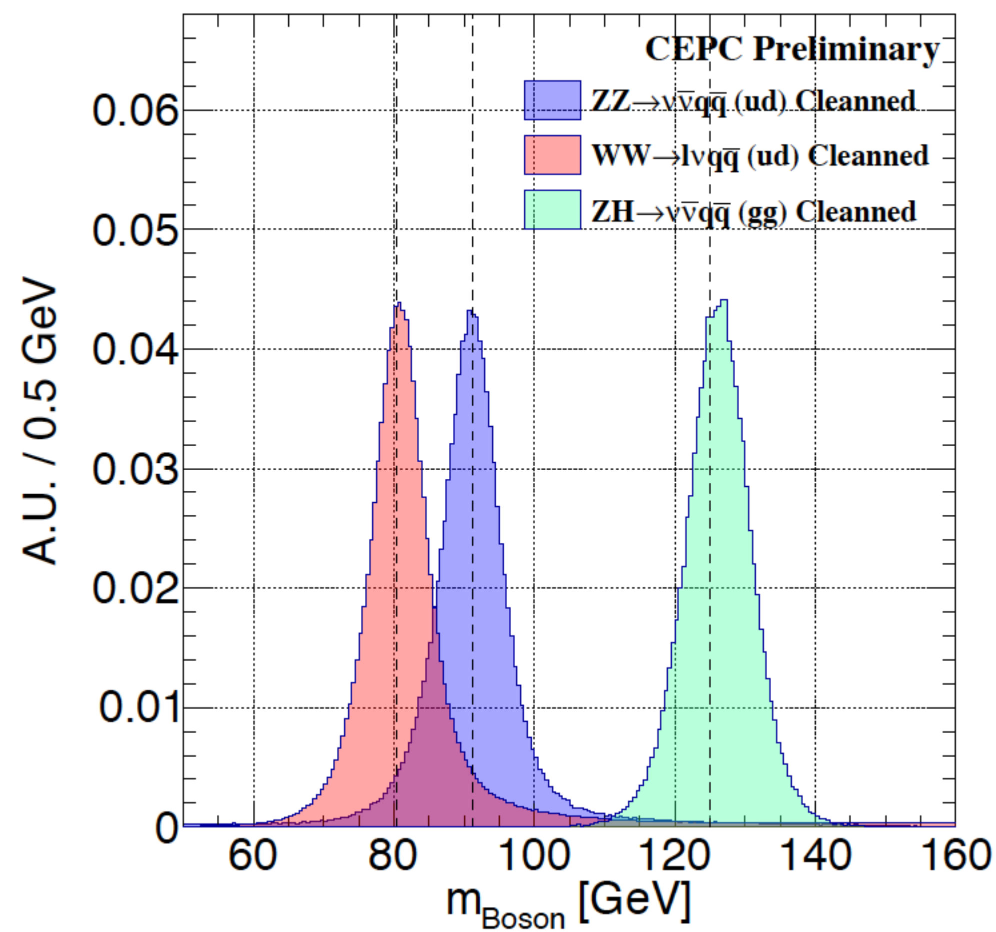
- Boson Mass resolution (BMR):
  - BM – 1: W, Z, H separation with hadronic final state
  - BM – 2: H→tautau measurement @ qqH

- BMR + Color Singlet identification (CSI)
  - BM – 3: Separation of WW&ZZ full hadronic event

- BMR + CSI + Flavor Tagging (FT)
  - BM – 4: H->bb, cc, gg measurement @ vvH/qqH

BM-I: Massive Boson Separation





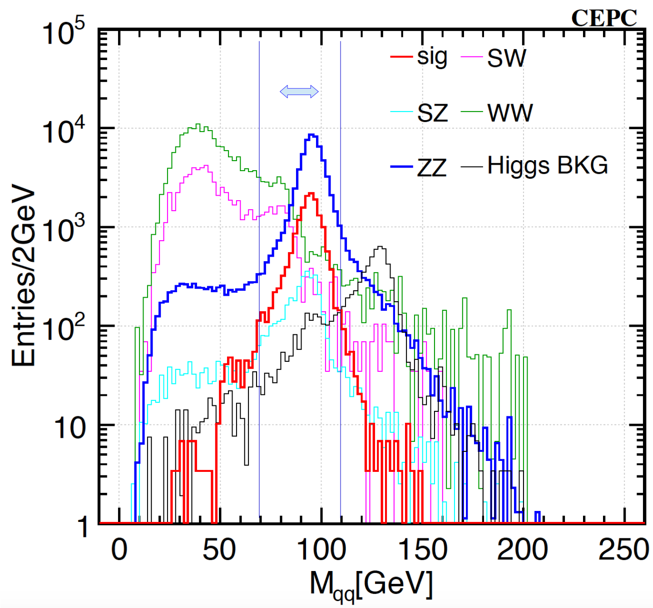
Peizhu Lai & CEPC CDR

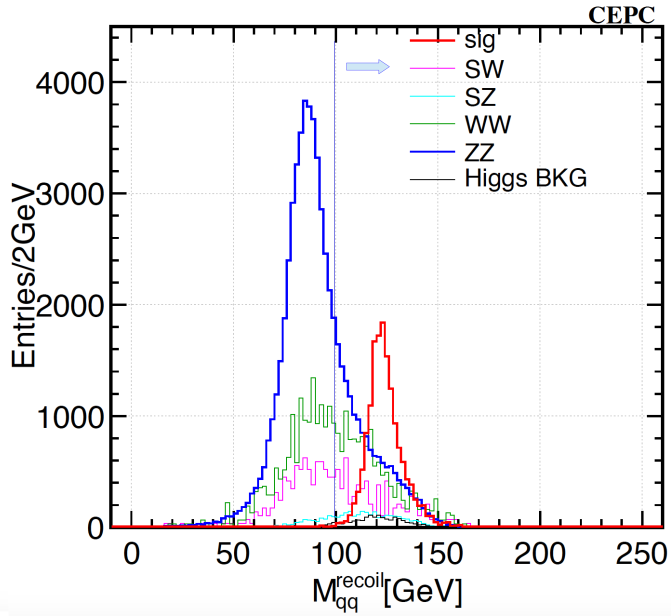
WW sample: using μvqq sample,
Plot: the visible mass without the muon

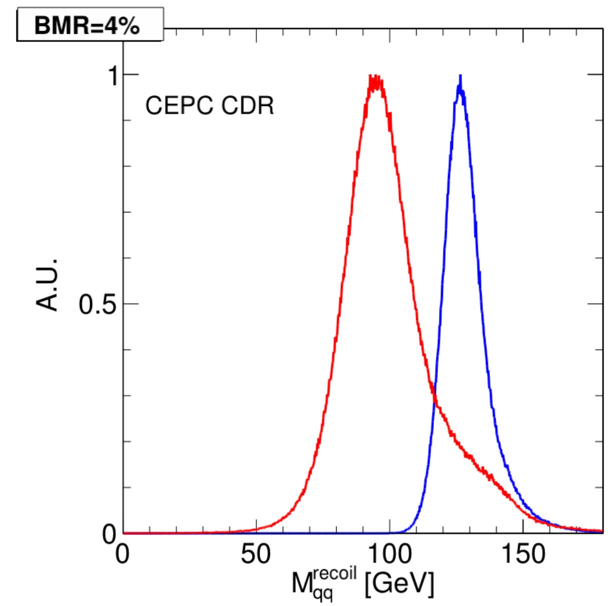
CEPC-RECO-2017-002 (DocDB id-164),
CEPC-RECO-2018-002 (DocDB id-171),

Eur. Phys. J. C (2018) 78: 426

BM-II: *g(Hττ)* at qqH



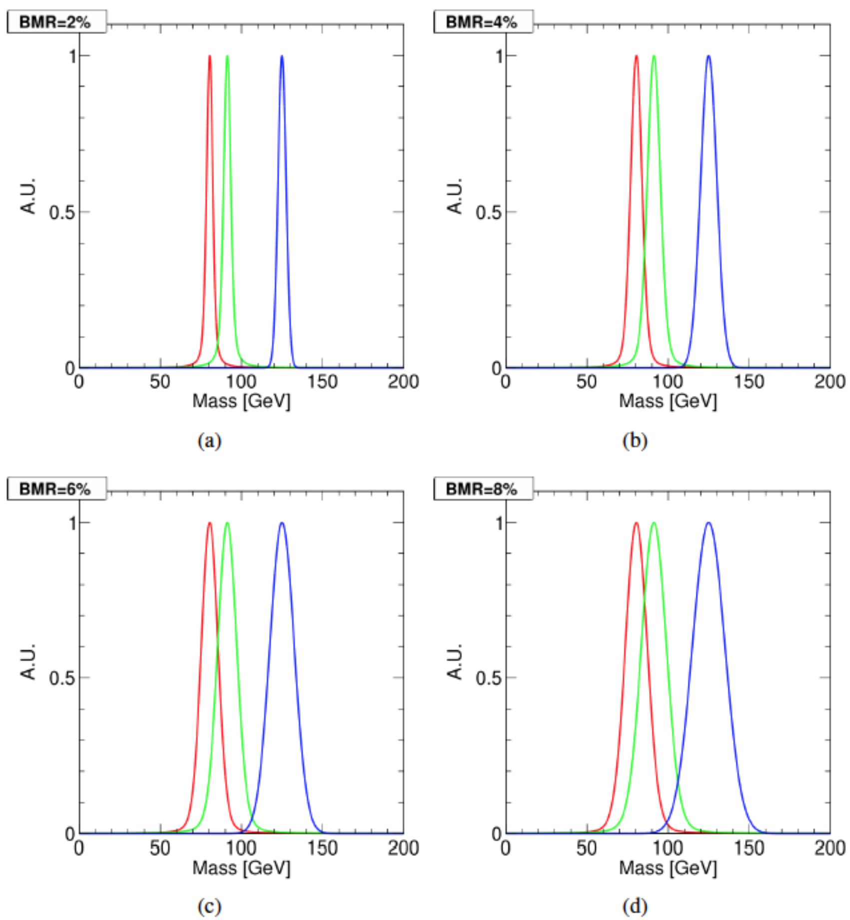


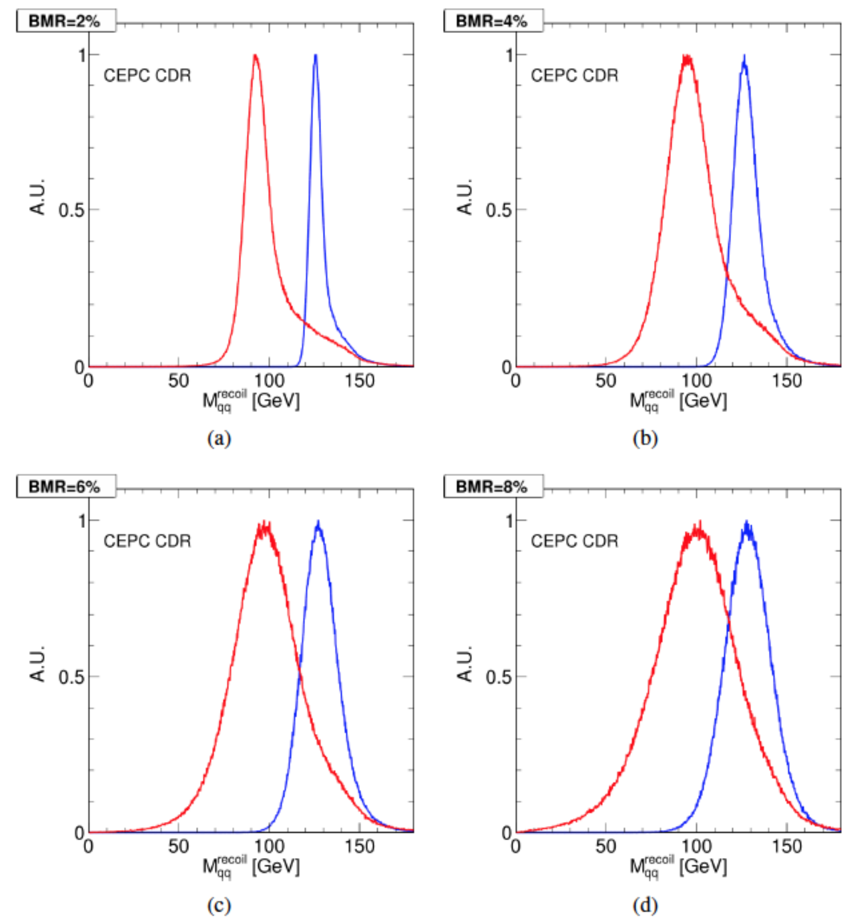


- TAURUS: di-tau system identification
- The rest particles are identified as the di-jet: to distinguish the ZZ/ZH background & Improves the accuracy by more than a factor of 2: **BMR < 4% (baseline of 3.8%) is crucial**
- Isolated tracks are intensionally defined as tau candidate: be distinguished by the VTX
- Relative accuracy of 0.9% at 5.6 ab^{-1} integrated luminosity, dominate the combined accuracy (0.8%)

Dan Yu's thesis Eur. Phys. J. C (2020) 80:7

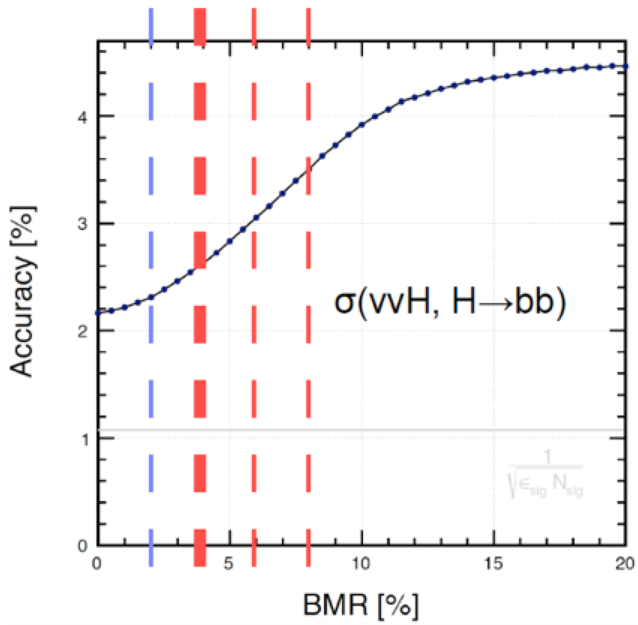
Key requirement: BMR < 4%

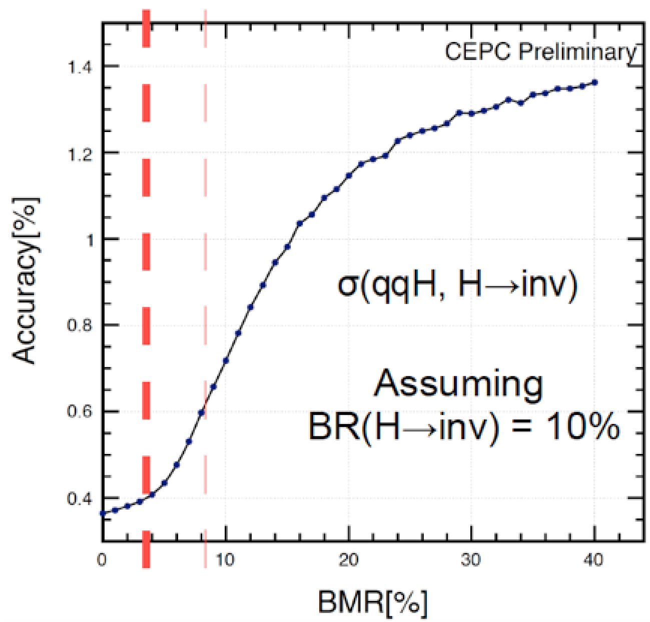


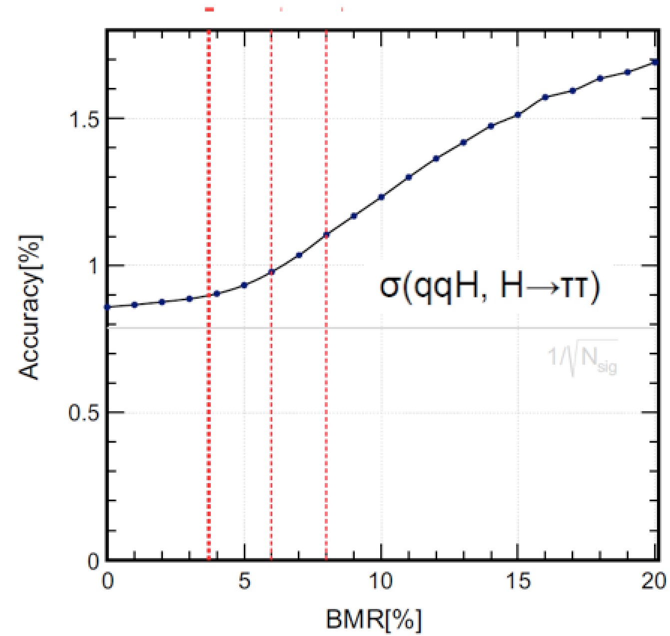


- BMR: relative mass resolution of the hadronic system, especially for the hadronically decayed massive Bosons: W, Z, H
- BMR < 4%: to separate qqH signal from qqX background with recoil mass

BMR V.S. benchmark accuracy



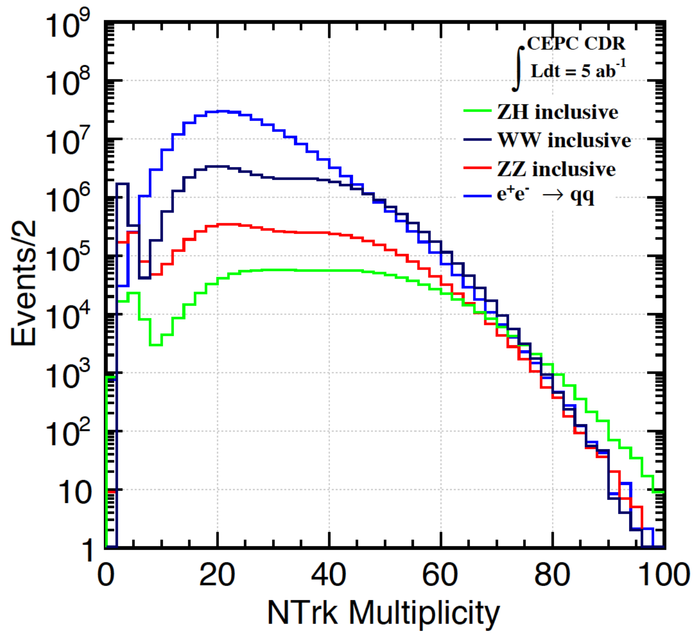


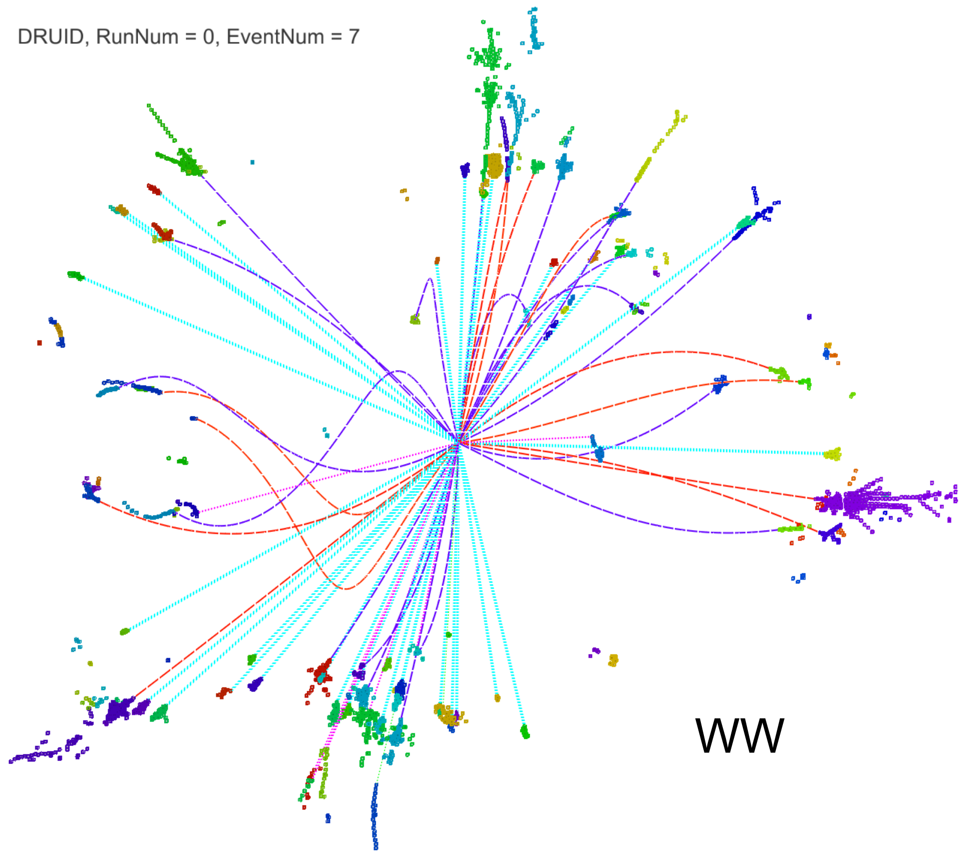


- Boson Mass Resolution: relative mass resolution of vvH, H→gg events
  - Free of Jet Clustering
  - Be applied directly to the Higgs analyses
- The CEPC baseline reaches 3.8%

	BMR = 2%	4%	6%	8%
σ(vvH, H→bb)	2.3%	2.6%	3.0%	3.4%
σ(vvH, H→inv)	0.38%	0.4%	0.5%	0.6%
σ(qqH, H→ττ)	0.85%	0.9%	1.0%	1.1%

BM-III: full hadronic WW-ZZ separation

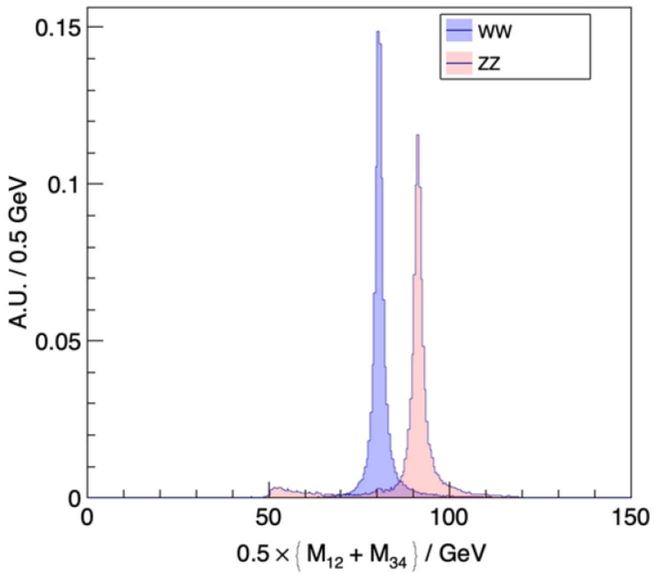


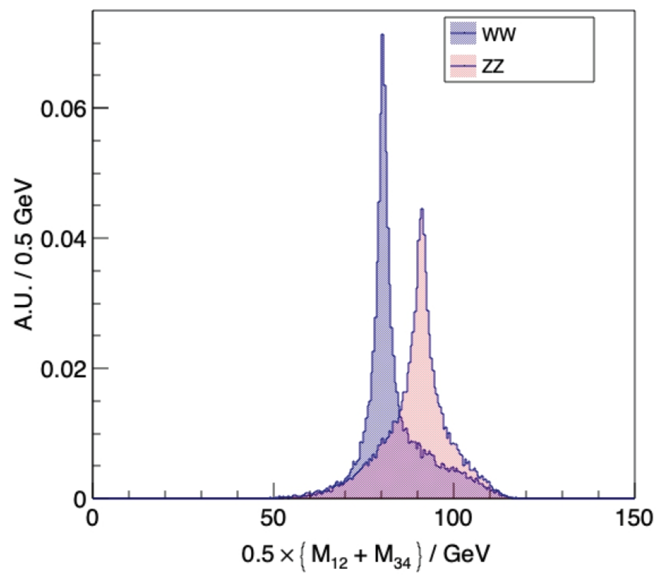


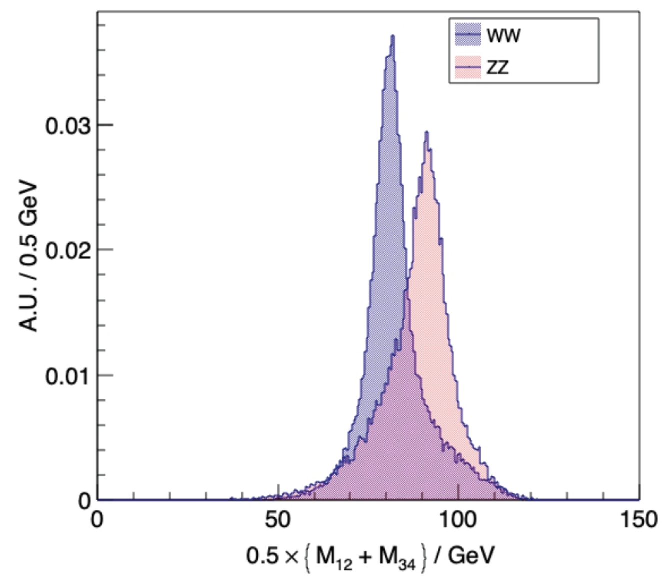
Eur. Phys. J. C (2019) 79:274

- Low energy jets! (20 – 120 GeV)
- Typical multiplicity ~ o(100)
- WW-ZZ Separation: determined by
  - Intrinsic boson mass/width
  - Jet confusion from color single reconstruction – jet clustering & pairing
  - Detector response

Jet confusion: the leading term







- Separation be characterized by

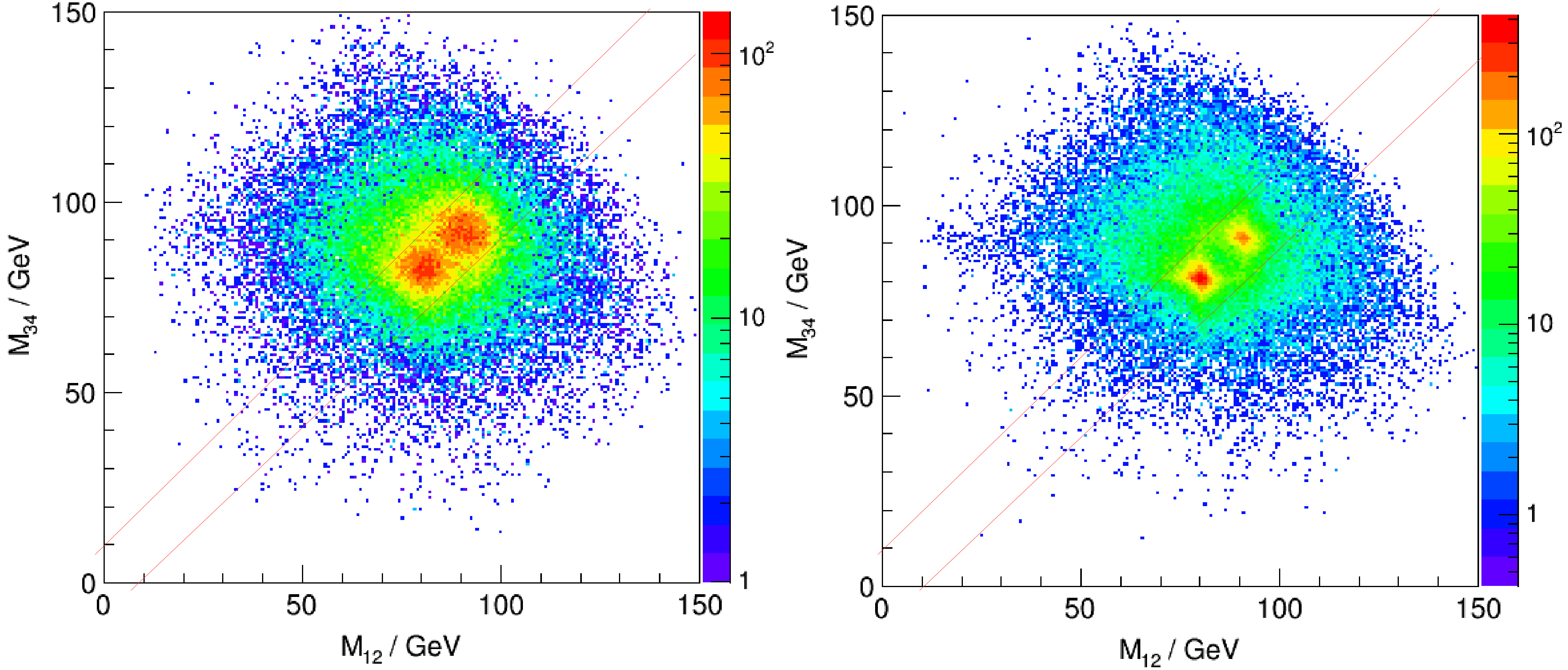
$$overlapping\ ratio = \sum_{bins} min(a_i, b_i)$$

- Final state/MC particles are clustered into Reco/Genjet with ee-kt, and paired according to chi2

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

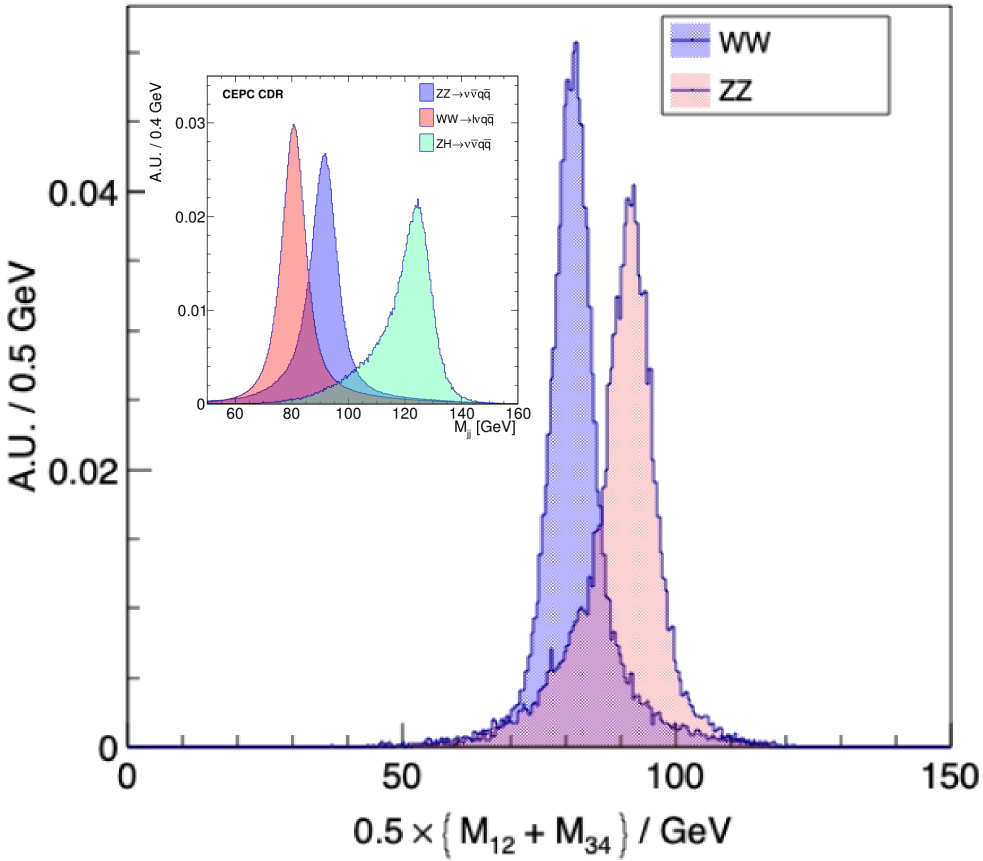
- WW-ZZ Separation at the inclusive sample:
  - Intrinsic boson mass/width - lower limit: Overlapping ratio of 13%
  - \+ Jet confusion – Genjet: Overlapping ratio of **53%**
  - \+ Detector response – Recojet: Overlapping ratio of 58%

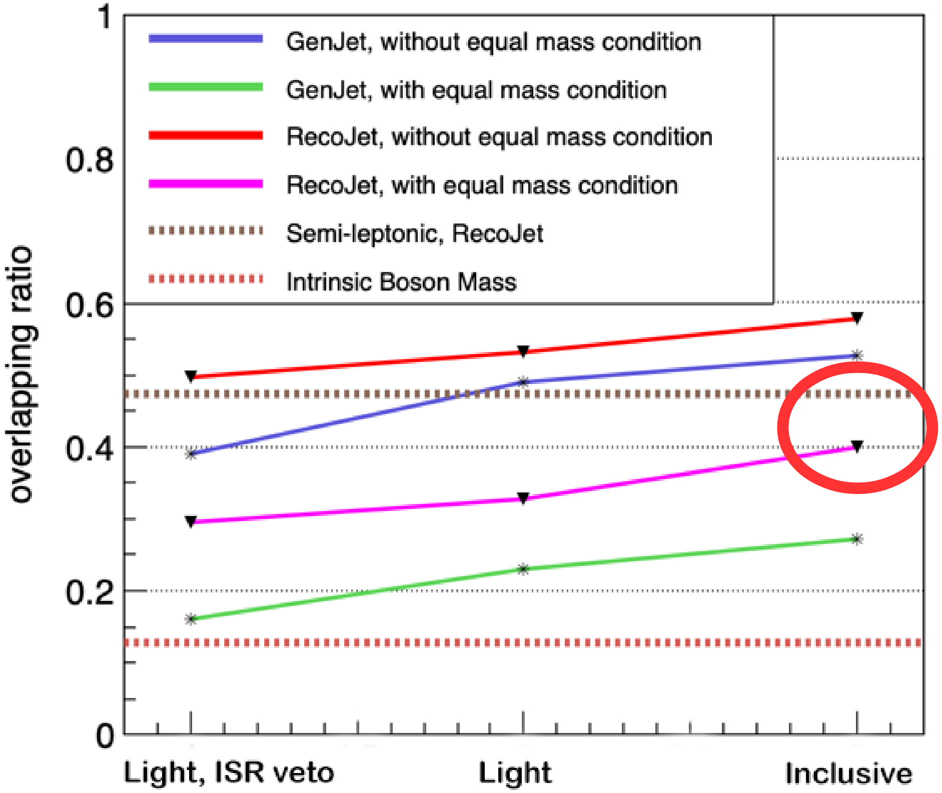
Reconstructed mass of the two di-jet system



Equal mass condition |M12 – M34| < 10 GeV: At the cost of half the statistic, the overlapping ratio can be reduced from 58%/53% to 40%/27% for the Reco/Genjet

Separation of full hadronic WW-ZZ event





The CEPC Baseline could separate efficiently the WW-ZZ with full hadronic final state.

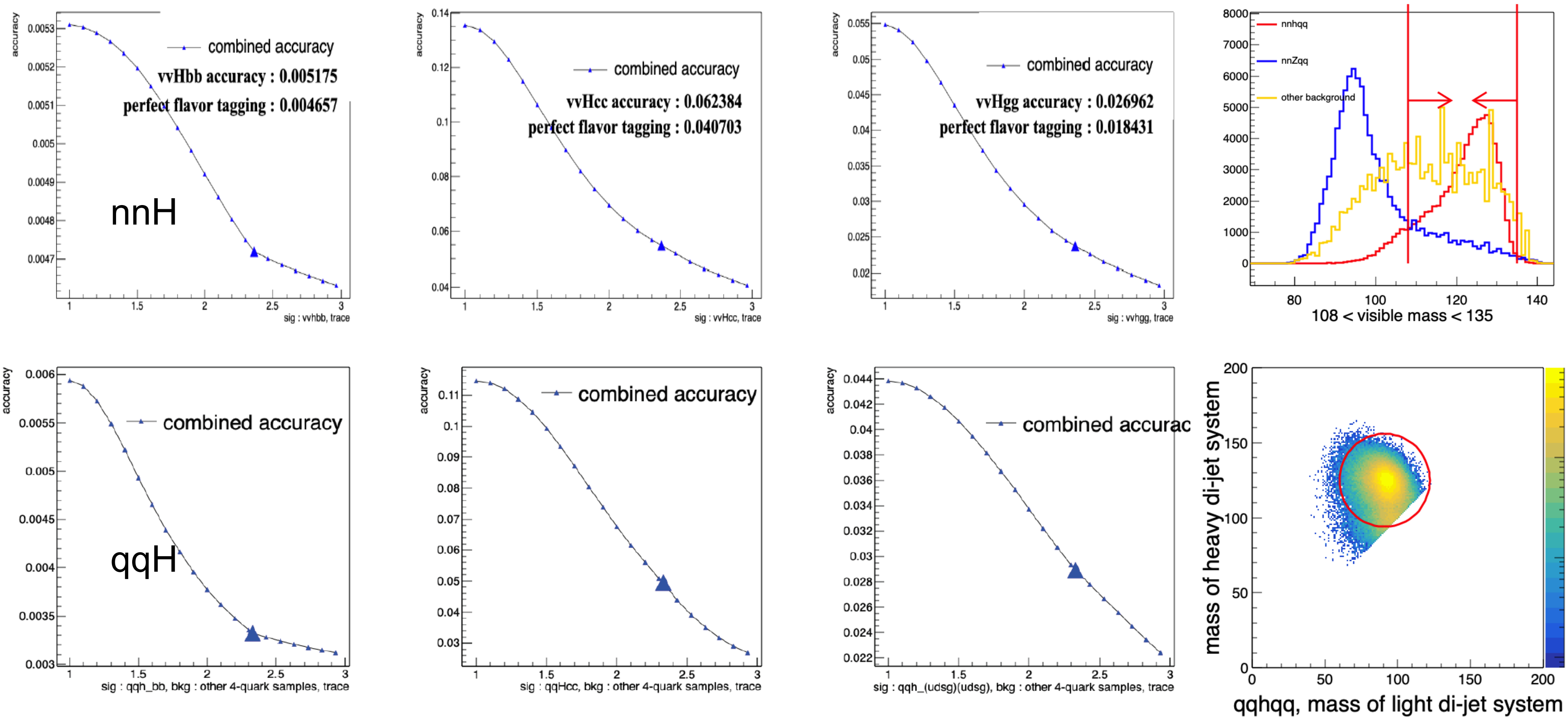
Critical to develop color singlet reconstruction: improve from the naive Jet clustering & pairing.

Quantified by differential overlapping ratio.

Control of ISR photon/neutrinos from heavy flavor jet is important.

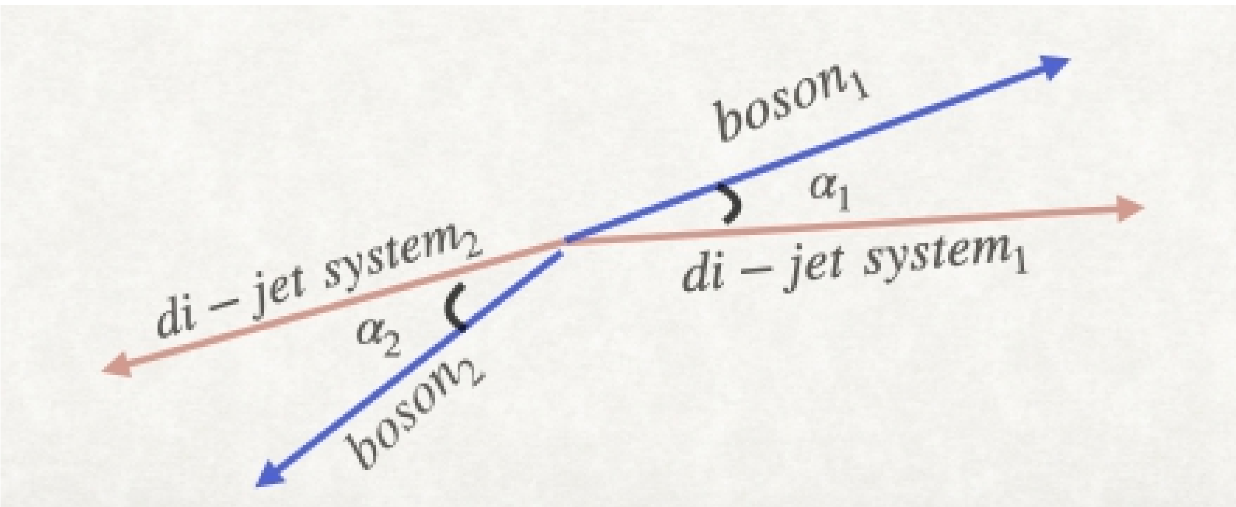
 https://arxiv.org/abs/1812.09478

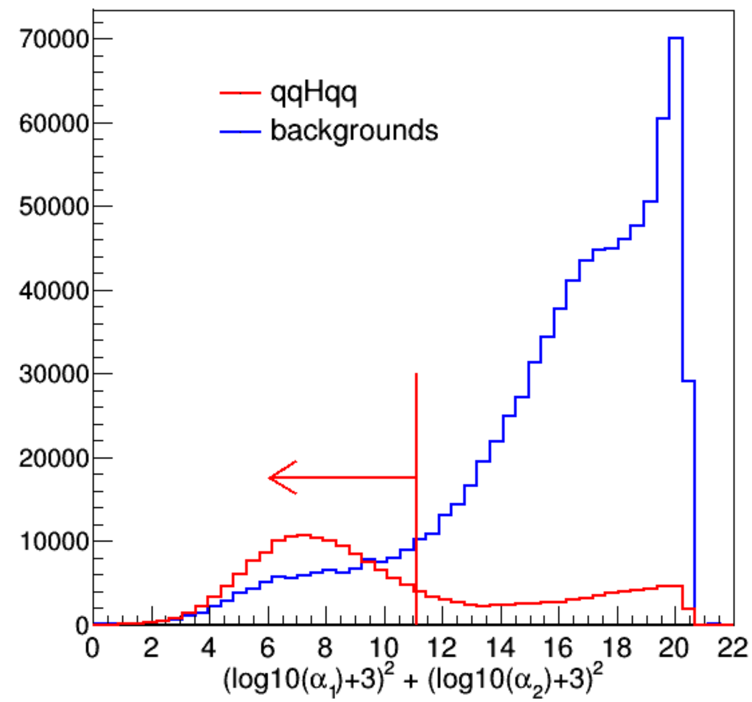
BM-IV: H→bb, cc, gg @ vvH - qqH



- BMR of 3.8% is good enough...
- Ideal Flavor tagging (Tr ~ 3) improves the accuracy of of Hcc by 2 times @ qqH, & 50% @ nnH

BM-IV: H→bb, cc, gg @ vvH - qqH





- CSI is still the bottle neck!
- CSI performance can be quantified by angles between truth/reco level bosons.
- A naive cut on angles improves the H→cc/gg by 50%/30%. See Yongfeng's talk at Thursday.

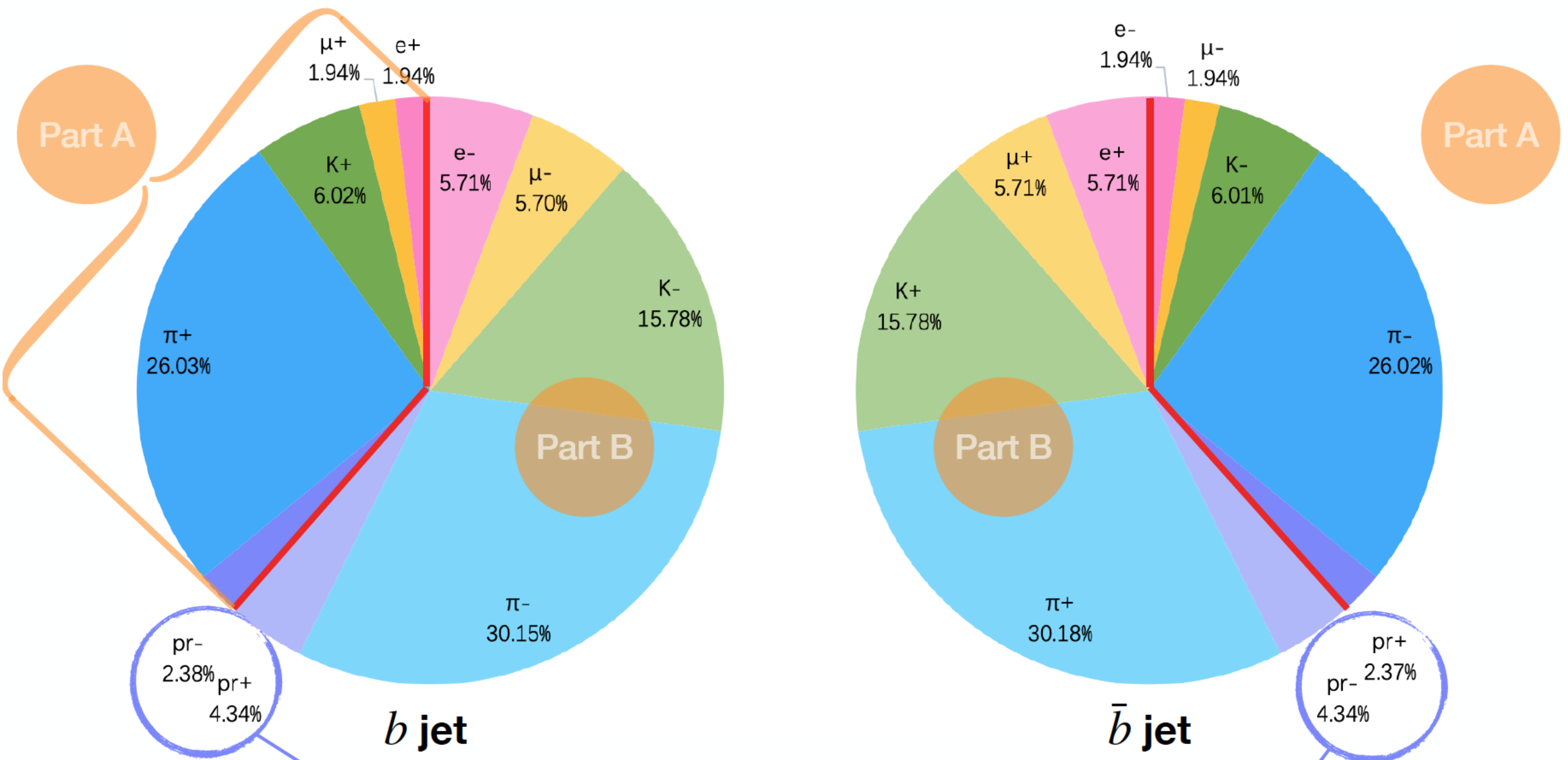
Summary

- Boson Mass Resolution: Better than 4% is critical, Baseline reaches 3.8%
- Jet energy scale & Jet energy resolution
  - Superb than LHC by 2-4 times @ JER, up to 1 order of magnitude @ JES. However, dependence to the jet clustering & matching need caution.
- Color Singlet Identification:
  - Bottleneck for measurements with more than 4-jets. Requires innovative developments/theoretical studies
- Jet Flavor Tagging:
  - Reasonable performance at CDR/baseline, further optimization on geometry/algorithm are critical for g(Hcc), g(Hbb), g(Hgg) measurements, etc.
- Jet Charge Reconstruction
  - Effective tagging power of 14%/28% @ $Z \to bb/cc$ achieved with straight forward algorithm (Hanhua's talk at performance session Nov 11th)
- Future: more benchmarks, especially differential analyses...

Backup

$Z \to b\bar{b}$

Percent of final charged leading particles



ω(using only charge) = 0.403

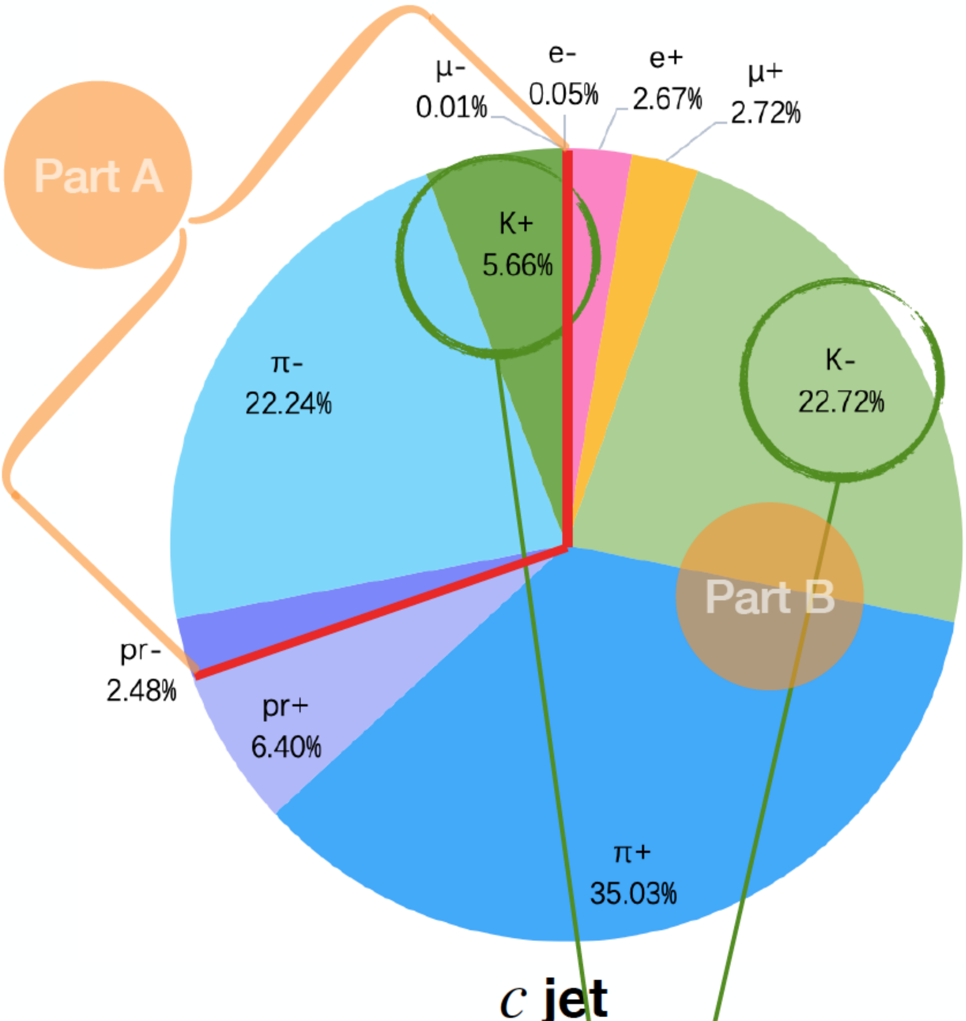
ω(using charge & PID) = 0.383

ω(using only charge) = 0.402

ω(using charge & PID) = 0.383

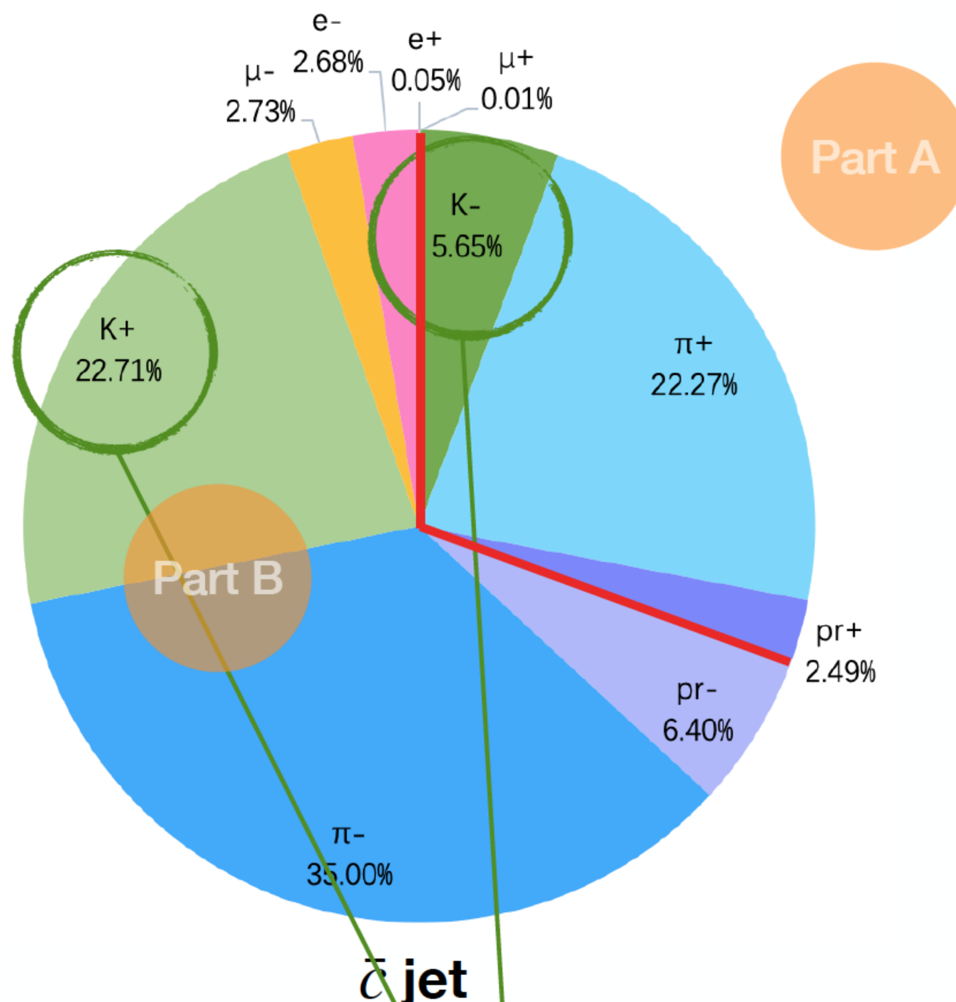
$Z \to c\bar{c}$

Percent of final charged leading particles



ω(using only charge) = 0.473

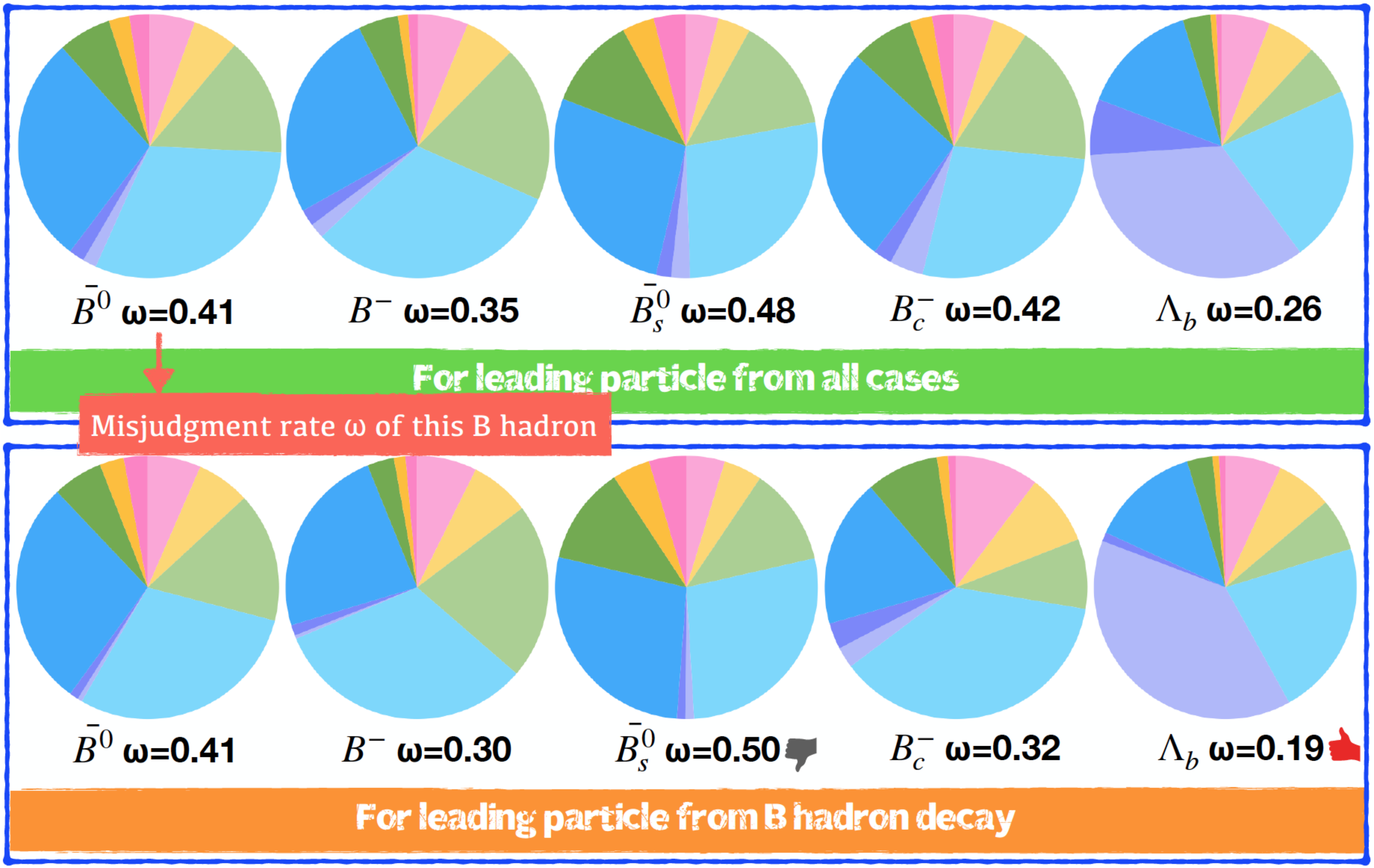
ω(using charge & PID) = 0.304



ω(using only charge) = 0.475

ω(using charge & PID) = 0.305

$Z \to b\bar{b}$ Percent of leading particles of each B hadron of b jet



To QCD community:

What's the dependency on fragmentation method/tools?

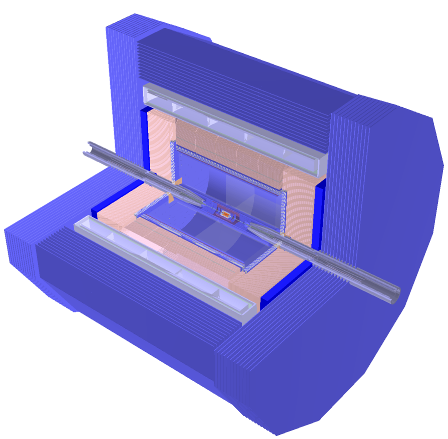
What's the ultimate CSI?

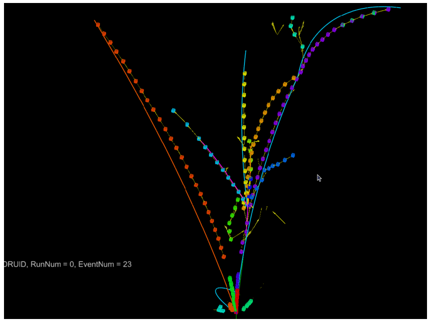
Summary

- The CEPC: extremely clean environment, well understood & adjustable initial states: opportunity & challenges for the QCD
- Multi-jets events
  - Great physics potential – needs more exploration
  - CSI is crucial & very interesting
  - Jet identification
- Baseline detector has promising performance, especially for the jets
  - Good understanding of the detector performance
  - Time to make impact on the detector design
- Pheno-Simulation, Looking forward to a firm & efficient collaboration
  - New observables & New methods

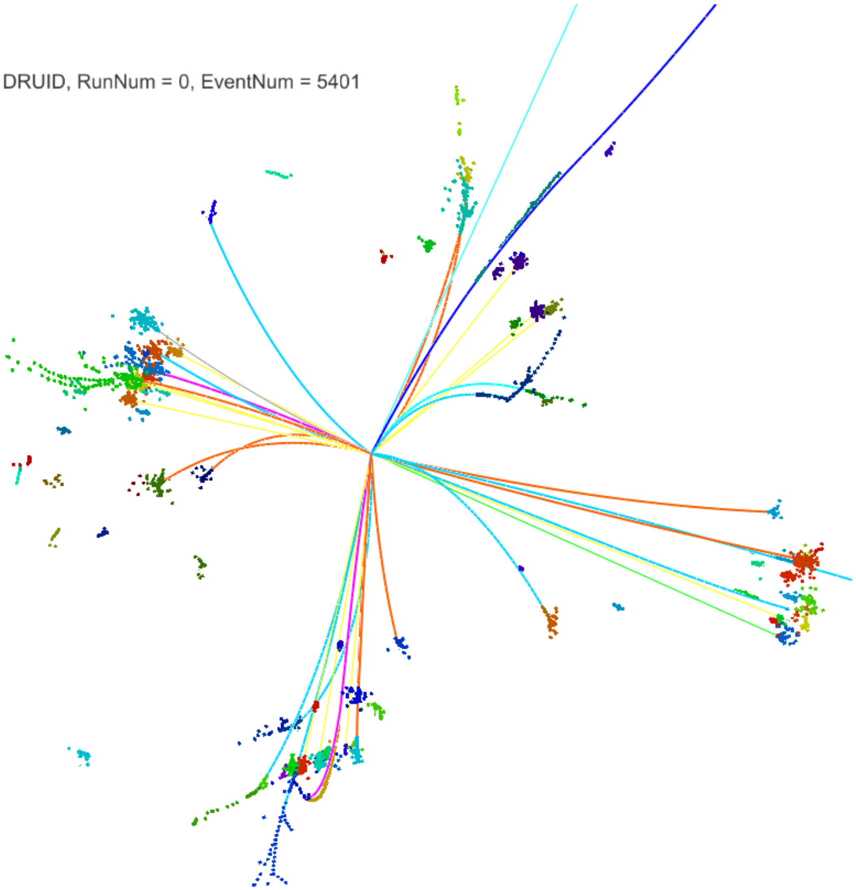
Baseline Detector Performance for flavor physics

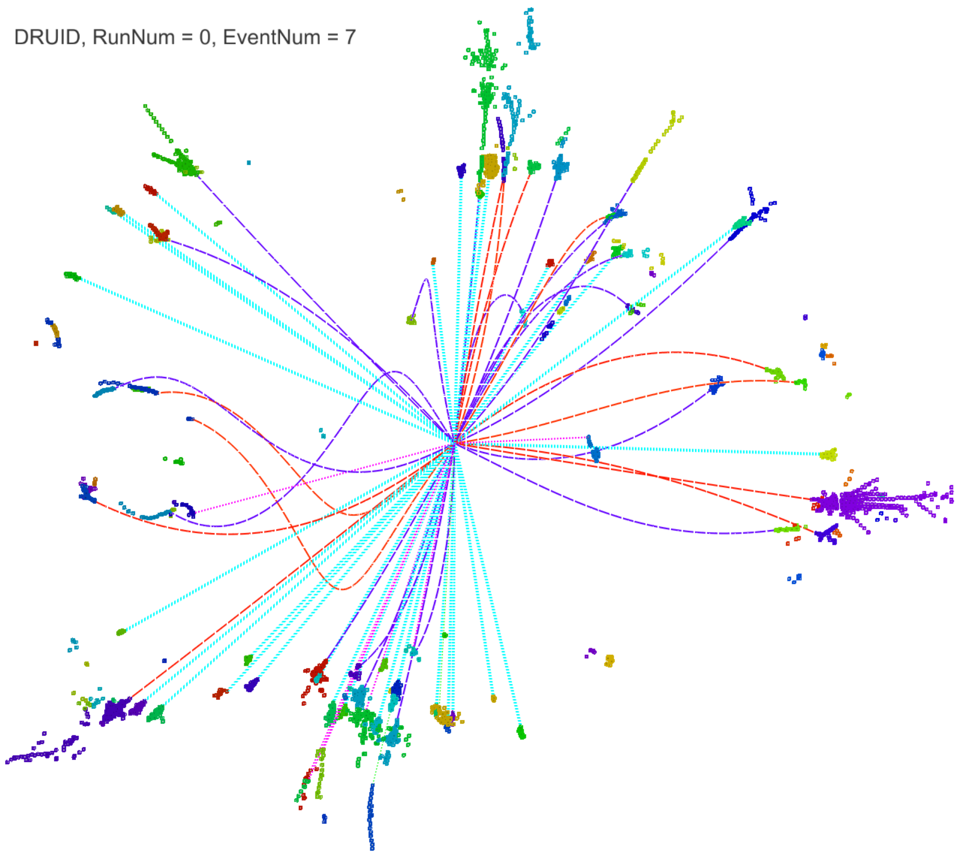
- Acceptance: $|\cos(\theta)| < 099$
- Tracks:
  - Pt threshold, ~ 100 MeV
  - $\delta p/p$ ~ o(0.1%)
- Photons:
  - Energy threshold, ~ 100 MeV
  - $\delta E/E$: 3 – 15%/sqrt(E)
- Pi-Kaon separation: 3-sigma
- Pi-0: rec. eff*purity @ $Z \to qq$ > 60% @ 5GeV
- B-tagging: eff*purity @ $Z \to qq$: 70%
- C-tagging: eff*purity @ $Z \to qq$: 40%
- Jet charge: eff*$(1-2\omega)^2$ ~ 15%/30% @ $Z \to bb/cc$
- Lepton inside jets: eff*purity @ $Z \to qq$ ~ 90% (energy > 3 GeV)
- Tau: eff*purity @ $WW \to tau\nu qq$: 70%, mis id from jet fragments ~ o(1%)
- Reconstruction of simple combinations: Ks/Lambda/D with all tracks @ $Z \to qq$: 60/75 – 80/85%
- BMR: 3.7%
- Missing Energy: Consistent with BMR.





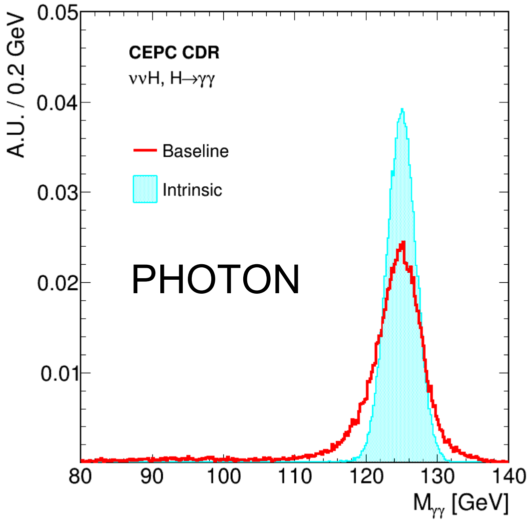
Color Singlet Identification...



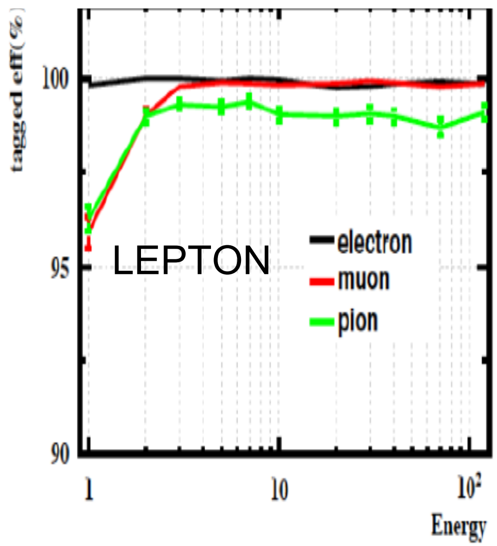


Identify the final state particles corresponding to one Boson decay.

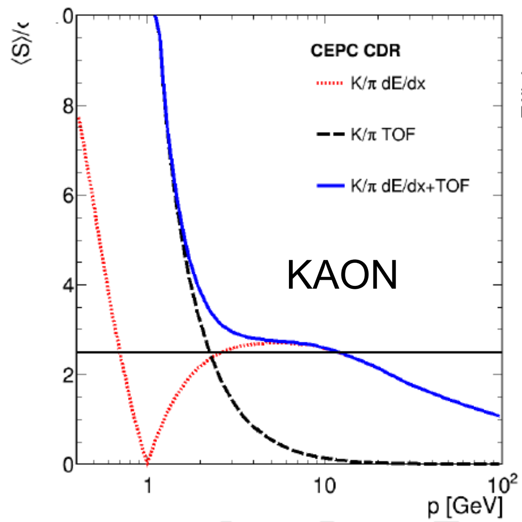
Physics Objects

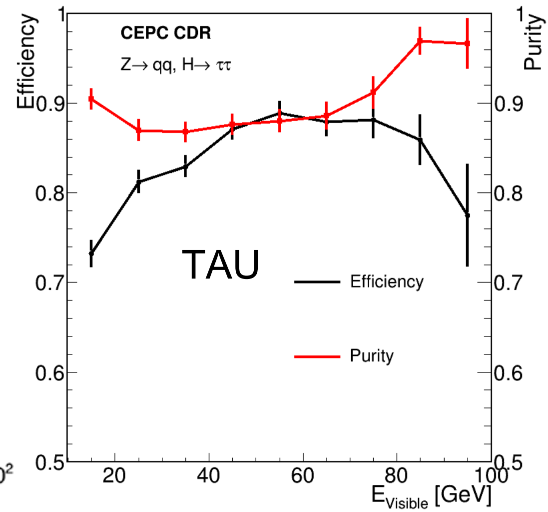


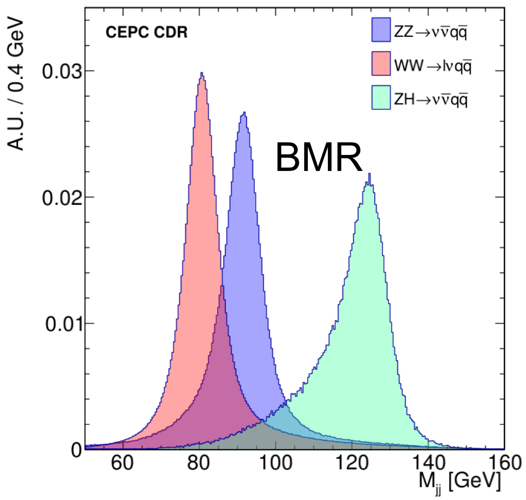
Eur. Phys. J. C (2017) 77: 591

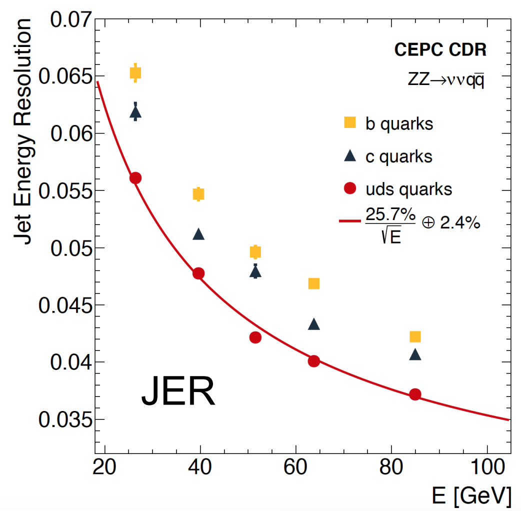


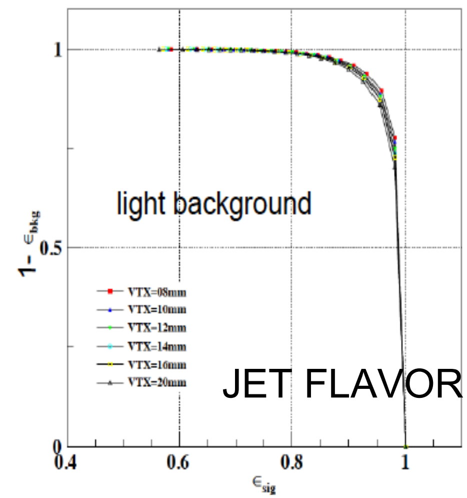
Eur. Phys. J. C (2018) 78:464

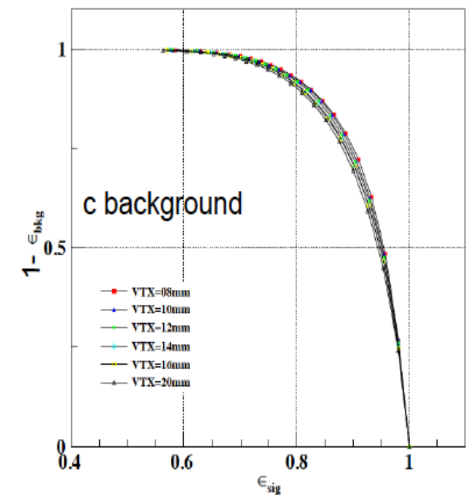






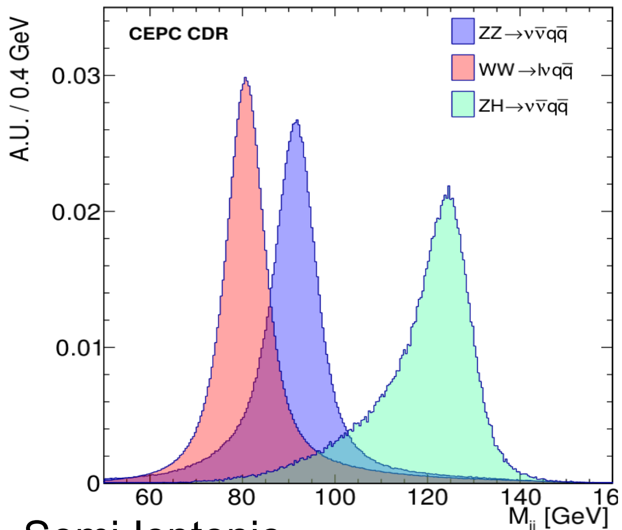


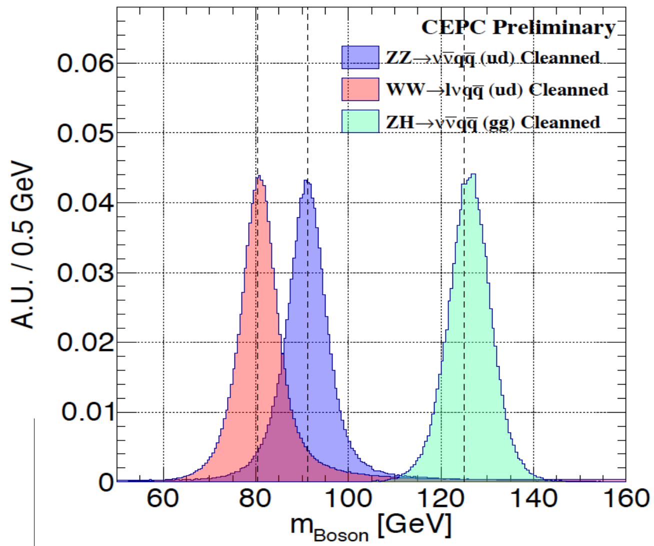




Eur. Phys. J. C (2018) 78: 426

Boson Separation in hadronic final states





Clear separation at inclusive Samples.
Could be significantly improved using cleaning

Eur. Phys. J. C (2018) 78: 426

Eur. Phys. J. C (2019) Submitted

Semi-leptonic

Full hadronic

Inclusive cleaned