CEPC Workshop 2026

Probing the CP structure of the Higgs–gluon coupling

Decay-plane correlations in H→gg at CEPC: tagging, categorization and first reconstruction results

Shahd Younes · IHEP

Outline

1. Motivation
2. Simulation workflow
3. Jet tagging
4. Event categorization

1. Decay-plane angle φ
2. Truth-level φ
3. Reconstructed φ and CP reach
4. Summary and next steps

Motivation: spin correlations in H→gg

The Hgg coupling can carry a CP-mixing phase Δ, with a CP-even term cosΔ and a CP-odd term sinΔ.

The two gluons are produced with equal helicity, so the decay planes of their splittings are correlated: 1 + A cos(2φ − 2Δ).

Wang, Gao, Li and Zhu estimate that qqqq events could probe Δ ≈ 0.03π at 20 ab⁻¹ with their E4C observable (n = 4), assuming ideal splitting-mode identification. gggg is much weaker. Our φ is Lund-like (z > 0.1), which is less sensitive in their study.

Our question: how much of this survives showering, reconstruction and tagging?

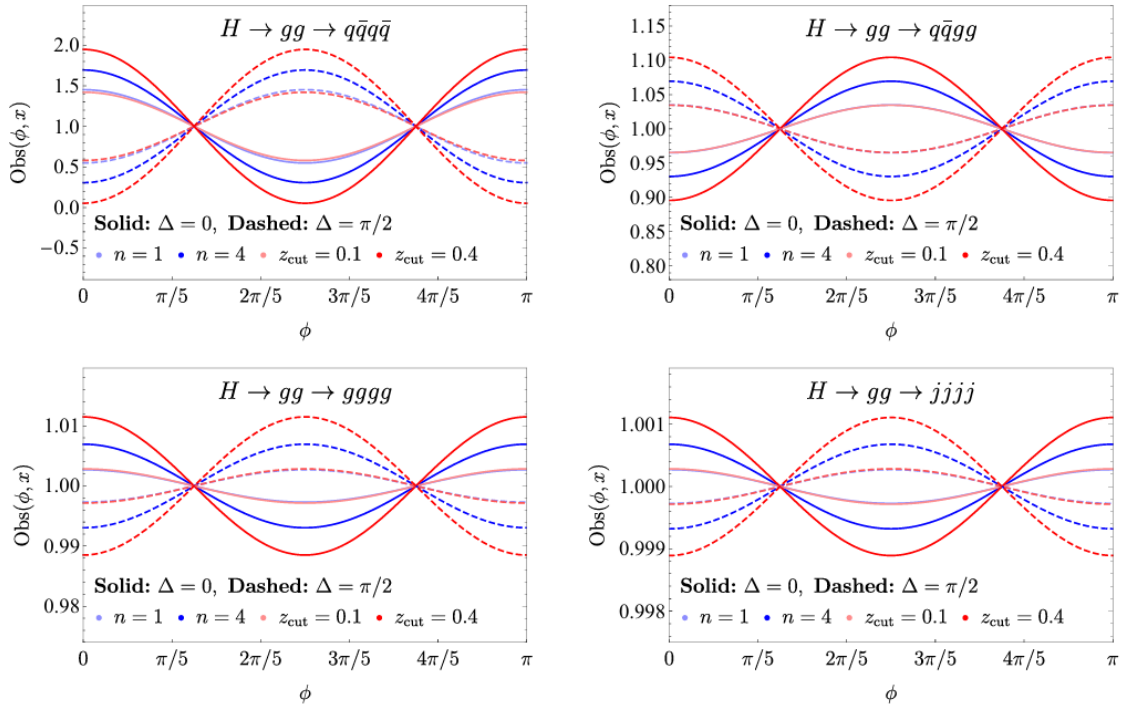


FIG. 3. $\mathrm{Obs}(\phi, x)$ for the Lund observable with $z_{\mathrm{cut}} = 0.1$ and $z_{\mathrm{cut}} = 0.4$, as well as for the E4C observable with $n = 1$ and $n = 4$. Results for both $\Delta = 0$ ($\mathcal{CP}$-even) and $\Delta = \pi/2$ ($\mathcal{CP}$-odd) in different splitting mode are presented. Since $A_{\mathrm{E4C}}(n = 1)$ and $A_{\mathrm{Lund}}(z_{\mathrm{cut}} = 0.1)$ are very close, the corresponding curves overlap.

Wang et al., arXiv:2601.22248, Fig. 3: solid Δ = 0 (CP-even), dashed Δ = π/2 (CP-odd).

Simulation workflow

1

Whizard 3.1.7

$e^+e^- \to ZH$ at 240 GeV, $Z \to \mu\mu$. Higgs decays from the HC_UFO_QCD model with CP-even or CP-odd couplings.

2

Pythia 8.310

Parton shower and hadronisation, linked to Whizard.

3

CEPCSW

TDR version of the software: detector simulation and particle-flow reconstruction, with the Z-muon fix.

4

Jets and ntuples

Z muons removed, two exclusive jets, one row per jet with up to 50 particles.

5

Analysis

Decay-plane angle φ from the sub-jets.

Truth samples: H→gg→gggg, ggqq and qqqq, generated separately for each CP hypothesis, with E > 5 GeV and m(ij) > 10 GeV. The quarks are u and d.

Standard Model sample: 100 seeds of 2000 events. 185,255 events pass the Z-muon pair and gluon-label requirements for tagging, and the φ analysis uses 193,441.

Datasets used for the tagger

Six Higgs-decay samples. Hgg provides two classes, g_gg and g_qq, so the tagger has seven classes.

Class	Process	Generator	Files	Jets
Hbb	ZH, H→b$\bar{b}$	Whizard3+Pythia8	299	~150k
Hcc	ZH, H→c$\bar{c}$	Whizard3+Pythia8	299	~150k
Hss	ZH, H→s$\bar{s}$	Whizard3+Pythia8	292	~146k
Huu	ZH, H→u$\bar{u}$	Whizard3+Pythia8	297	~149k
Hdd	ZH, H→d$\bar{d}$	Whizard3+Pythia8	299	~150k
Hgg	ZH, H→gg	Whizard3+Pythia8	300	~150k

Jet tagger: ParticleTransformer

Input

Particles

Up to 50 particles per jet, 20 features each plus the four-vector.

Features

Kinematics and detector

pT, relative energy, ΔR, Δη, Δφ, charge, d0 and z0 with errors, ECAL and HCAL energy and cluster sizes, particle ID flags.

Network

Transformer

8 layers, 8 heads, embedding [128, 512, 128], pair embedding [64, 64, 64].

Training

Weaver

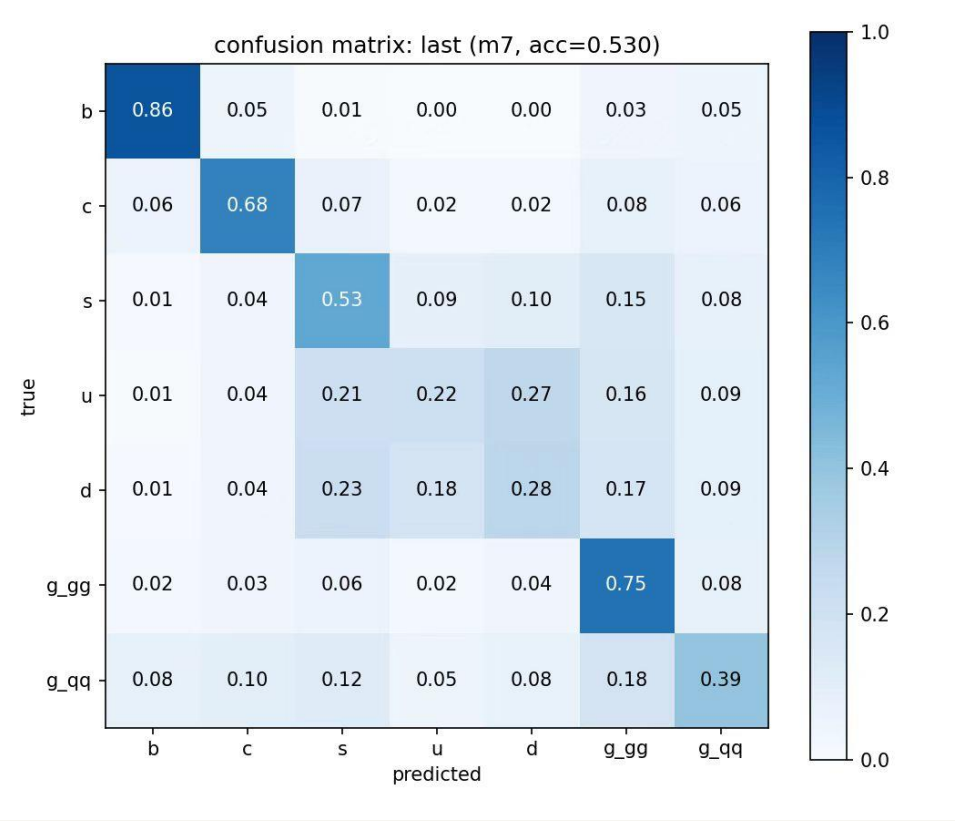
Ranger, learning rate 1e-3, batch 512, 50 epochs, separate train, validation and test files.

> Seven classes: b, c, s, u, d, g_gg and g_qq.

> The gluon label comes from the first identifiable splitting of each gluon, $g \to gg$ or $g \to q\bar{q}$, following the corrected generator-truth rule. The technique is the same as in the CEPC TDR.

Tagger results on the next two slides are from Y. Guo's m7 model. My own model is shown in the backup.

Tagger performance: g_gg versus g_qq



m7 model (Y. Guo): confusion matrix after the last epoch, normalised by true class.

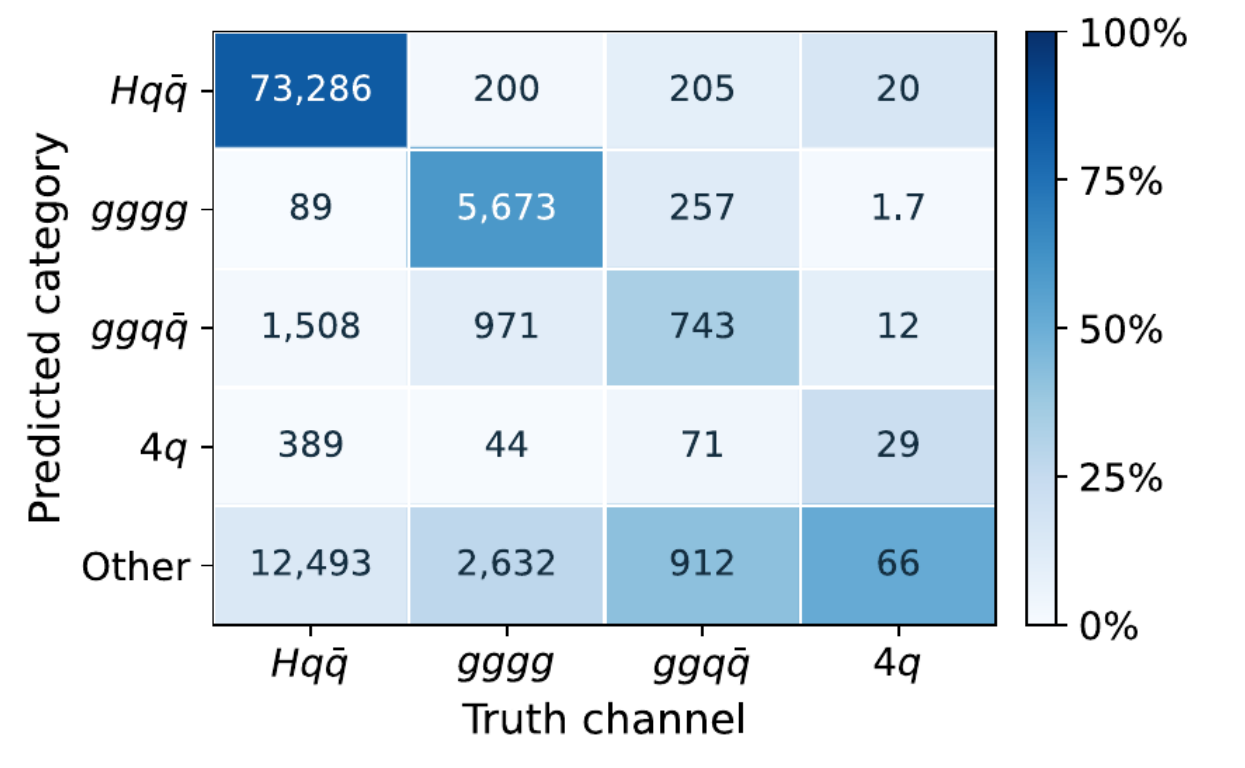
Accuracy	g_gg recall	g_qq recall
0.53	0.75	0.39

b and c jets are identified well. u, d and s overlap strongly.

For the gluons, 18% of true g_qq jets are called g_gg, and 8% of true g_gg jets are called g_qq.

Event categorization: expected yields

ZH, Z → μμ, 21.6 ab^{-1}. Hqq is H→bb̄, cc̄ and ss̄. Yields from Y. Guo's m7 model.



Expected events, rows: predicted category, columns: truth channel. "Other" (rule unverified) is excluded from the fit.

gggg region purity

94%

ggqq region purity

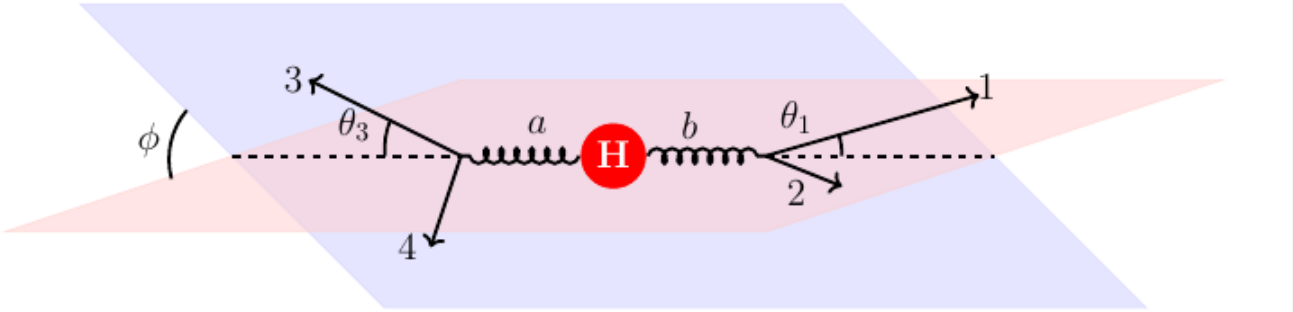
23%

qqqq region purity

5.4%

Mode fractions follow our generator-truth label (first identifiable splitting); the paper's one-step estimate is about 93% gggg, 7% ggqq and 0.1% qqqq.

Decay-plane angle φ



Wang et al., arXiv:2601.22248, Fig. 1.

Expected distribution

$$1 + A\cos(2\varphi - 2\Delta)$$

Δ = 0 for CP-even, Δ = π/2 for CP-odd.

1

Gluon splitting

Each gluon splits into two daughters: sub-jets in reco, partons at truth level.

2

Decay planes

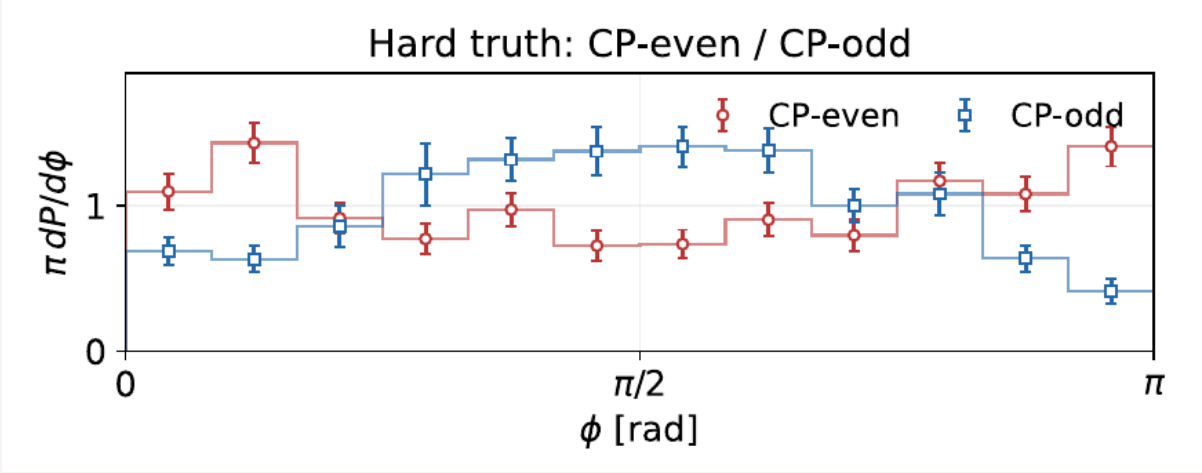
The two daughters of one gluon span a plane, so each event has two planes.

3

Angle φ

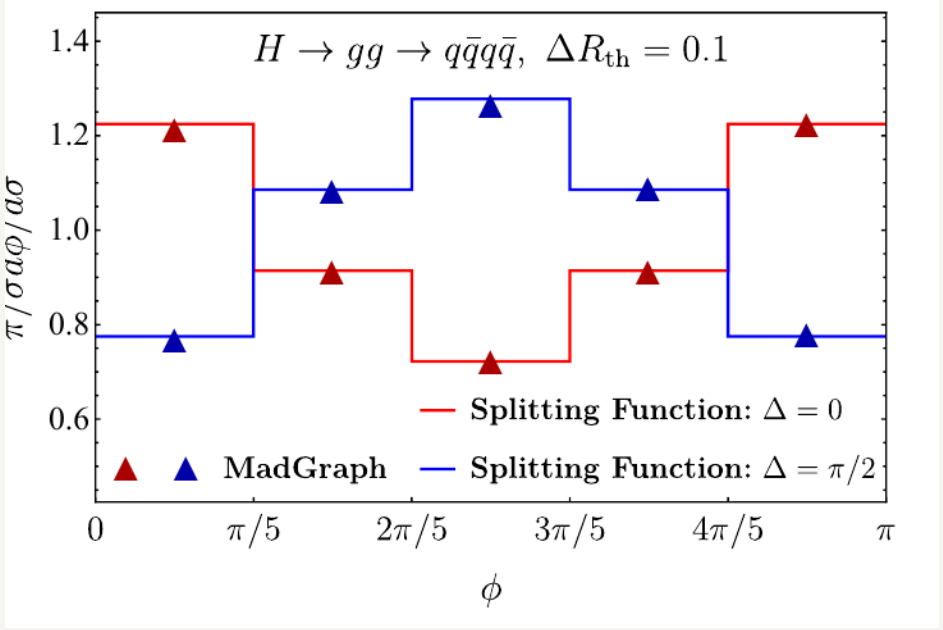
The angle between the planes, in the rest frame of the four daughters, mod π.

Truth-level φ for CP-even and CP-odd



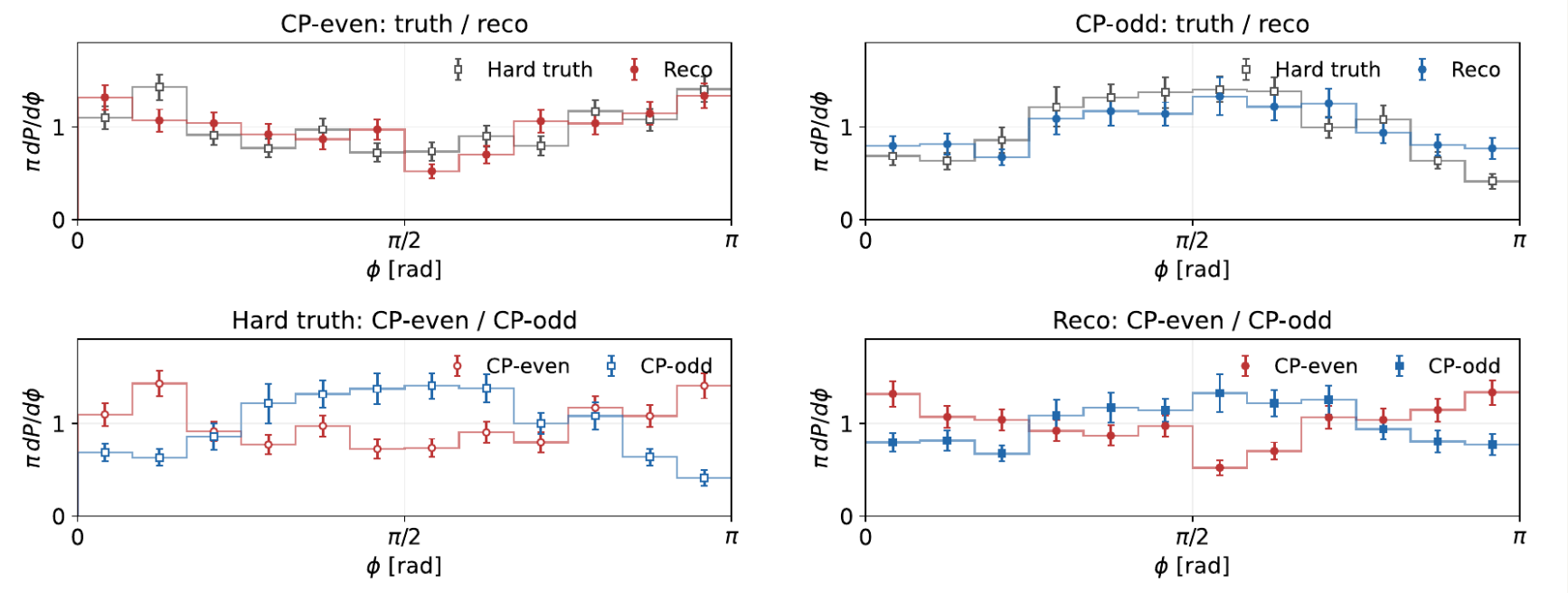
Four quarks, truth level: CP-even (red) and CP-odd (blue). The vertical axis is π dP/dφ, normalised to unit area.

The two hypotheses are mirror images, as in the paper. ggqq shows a small difference at parton level and gggg none.



Wang et al., arXiv:2601.22248, Fig. 2: red is Δ = 0 (CP-even), blue is Δ = π/2 (CP-odd).

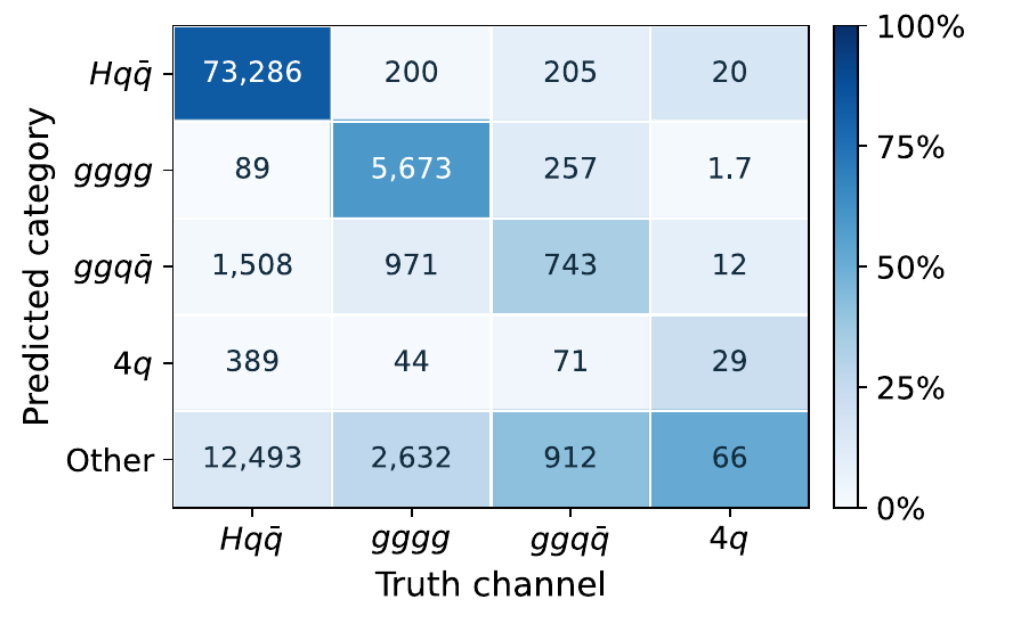
Reconstructed φ: CP separation survives, diluted



Four quarks: two exclusive e^+e^- kt jets, then Cambridge/Aachen declustering along the leading branch (max kt, $z > 0.1$), then φ and a frozen quality cut. Each plane uses the chosen split's daughters. 2177 of 13,800 events are selected, a weighted efficiency of 15.7%. The truth-to-reco dilution is 0.35 (backup), so the reco signal is weaker than at truth level.

240 GeV ZH, Z → μμ; HC + Pythia + CEPCSW (MC-assisted PID). Unit area, MC errors.

Expected sensitivity to the CP angle



Expected events at 21.6 ab⁻¹, Z → μμ. Rows: predicted category, columns: truth channel. Same matrix as slide 8. “Other” (rule unverified) is excluded from the fit.

Preliminary significance (truth-angle), CP-even true

	Categories	Ideal 4q
Reject CP-odd	0.54 σ	4.12 σ
Reject CP average*	0.27 σ	2.07 σ

Predicted 4q category: 5.4% purity. The ideal case uses all 129 true 4q events, with 100% purity and retention.

Preliminary: a truth-angle forecast, not a reco-level result. A more comprehensive analysis to probe CP mixing and spin correlations is planned. ⟨cos 2φ⟩ fit with ideal angles, fixed yields and flat Hqq, MC errors only. Assumes category–φ independence and HC to shower shape transfer; gggg is collinear. *CP average ≡ (even + odd)/2; equivalence to spin-off is untested.

Summary and next steps

- Tagging: the gggg category is 94% pure, but the ggqq and qqqq categories are only 23% and 5.4% pure.
- Truth level: qqqq shows a clean CP-even to CP-odd flip, ggqq a small one, gggg none.
- Reconstruction: for four quarks the CP separation survives but is diluted (0.35), with a 15.7% weighted selection efficiency.
- Preliminary reach at 21.6 ab^{-1} (truth-angle forecast): 0.54 σ to reject CP-odd with the categories, against 4.12 σ for ideal 4q. The reco-level reach will be lower.

Next steps

Event-level tagging

Our jet tagging applied at event level should significantly improve 4q purity (now 5.4%) and efficiency.

More statistics

Generate separate HC-model samples for gggg, ggqq and qqqq: more statistics for tagger training and analysis.

Raise the efficiency

Only 15.7% of four-quark events survive the φ selection. Event-level tagging should help here too.

Full CP analysis

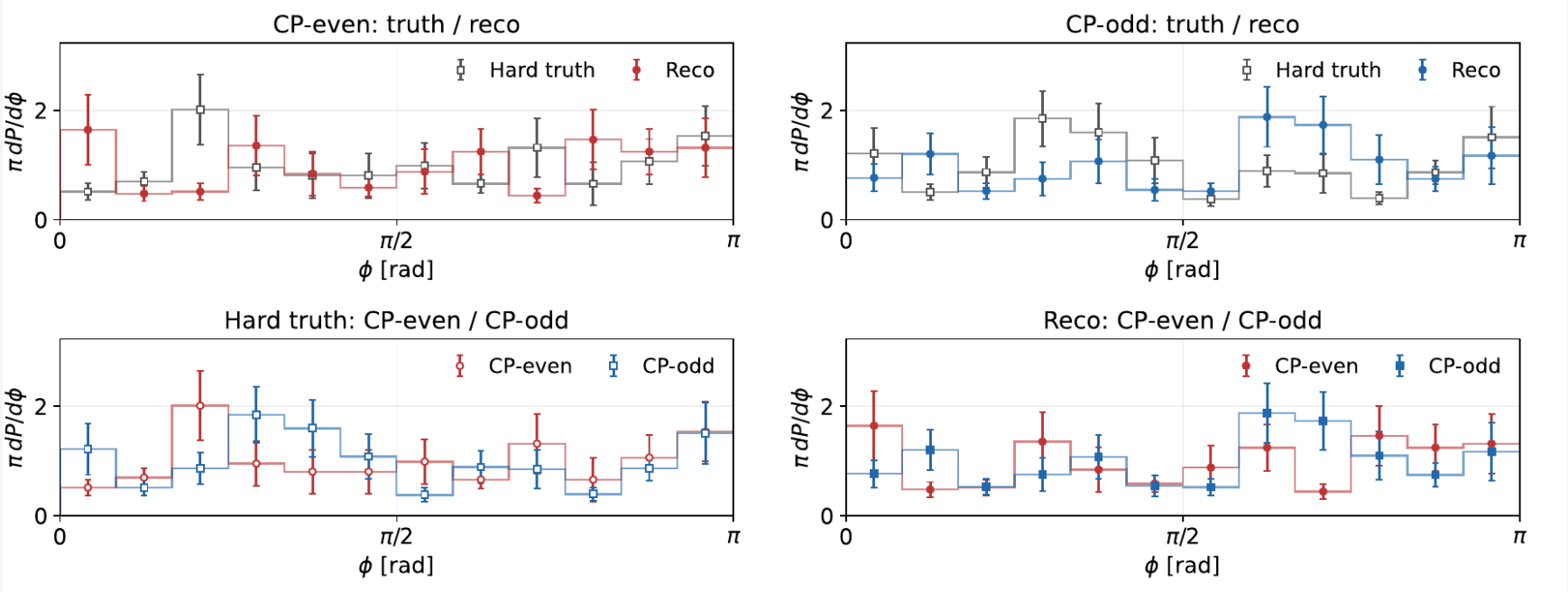
A more comprehensive analysis to probe CP mixing and spin correlations, fitting the reconstructed φ in the categories.

Thank you

Questions and comments are welcome.

Shahd Younes · IHEP · [email]

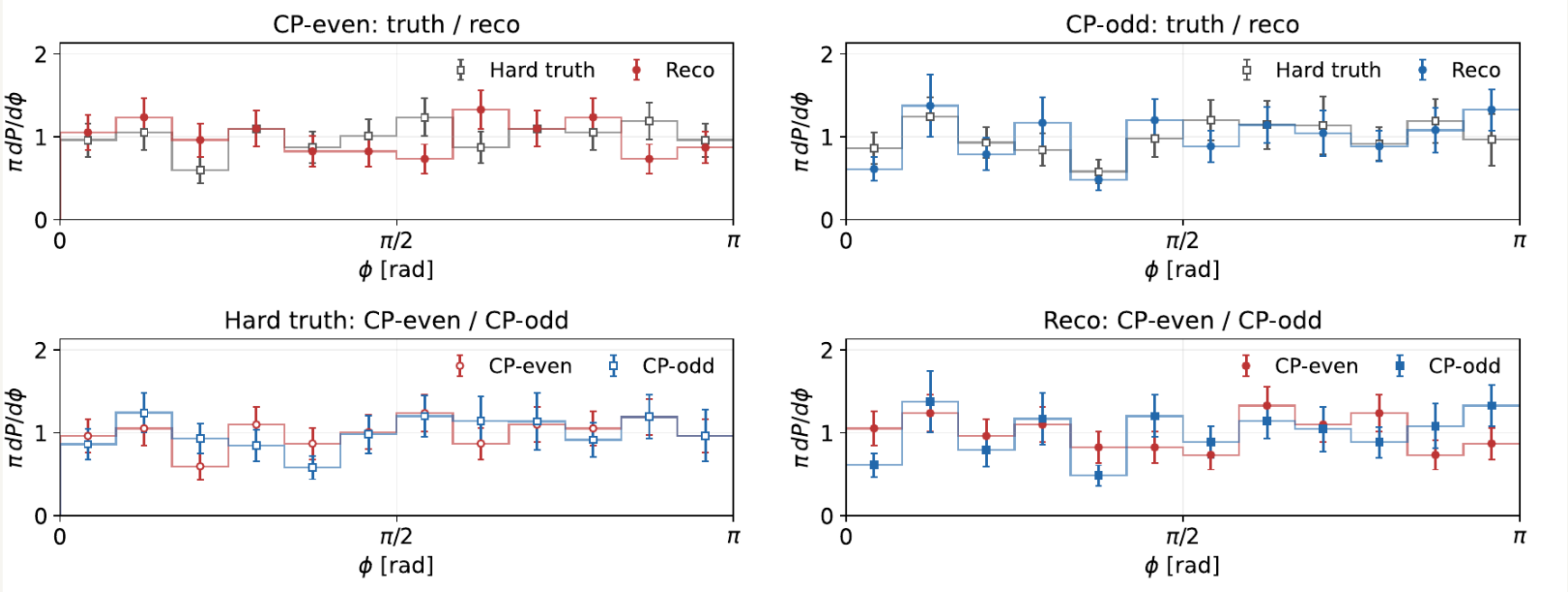
Backup: two gluons and two quarks, truth and reco



414 of 4,800 events selected, weighted efficiency 7.57%. The current sample does not resolve the reco CP difference.

Same reconstruction and frozen quality cut as for four quarks. 240 GeV ZH, Z → μμ; HC + Pythia + CEPCSW (MC-assisted PID).

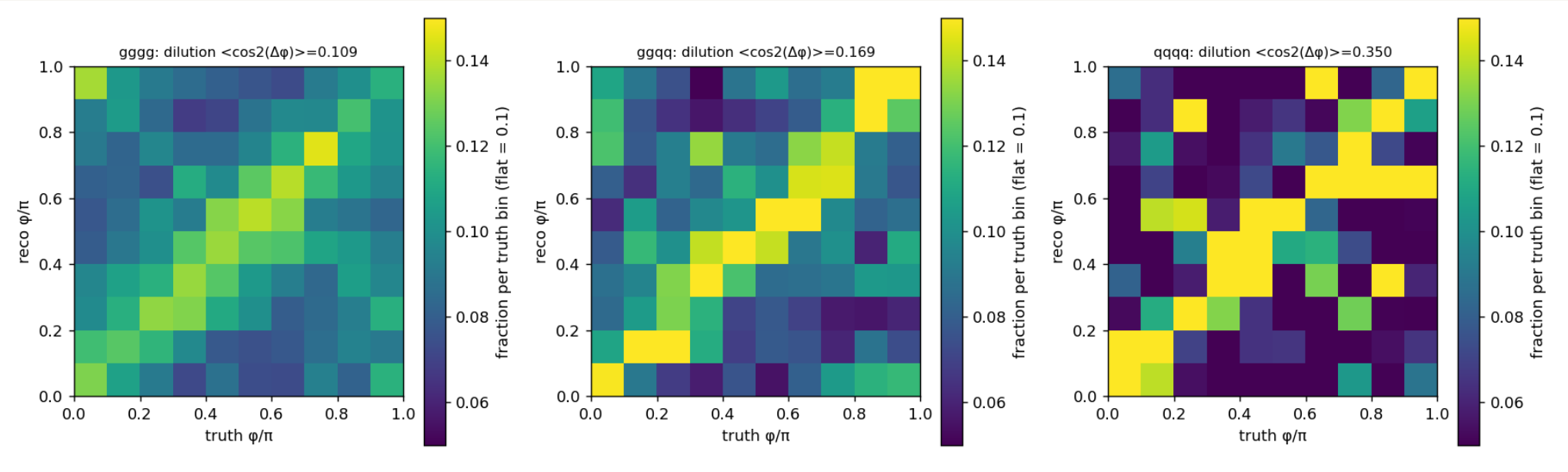
Backup: four gluons, truth and reco



560 of 12,400 events selected, weighted efficiency 4.45%. The current sample does not resolve the reco CP difference.

Same reconstruction and frozen quality cut as for four quarks. 240 GeV ZH, Z → μμ; HC + Pythia + CEPCSW (MC-assisted PID).

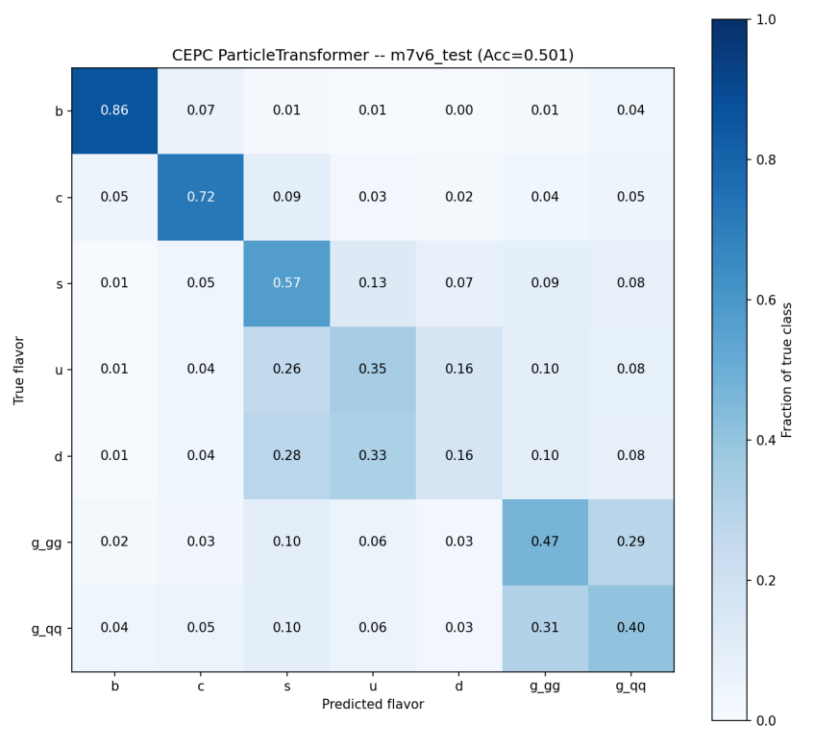
Backup: how much of the φ information survives

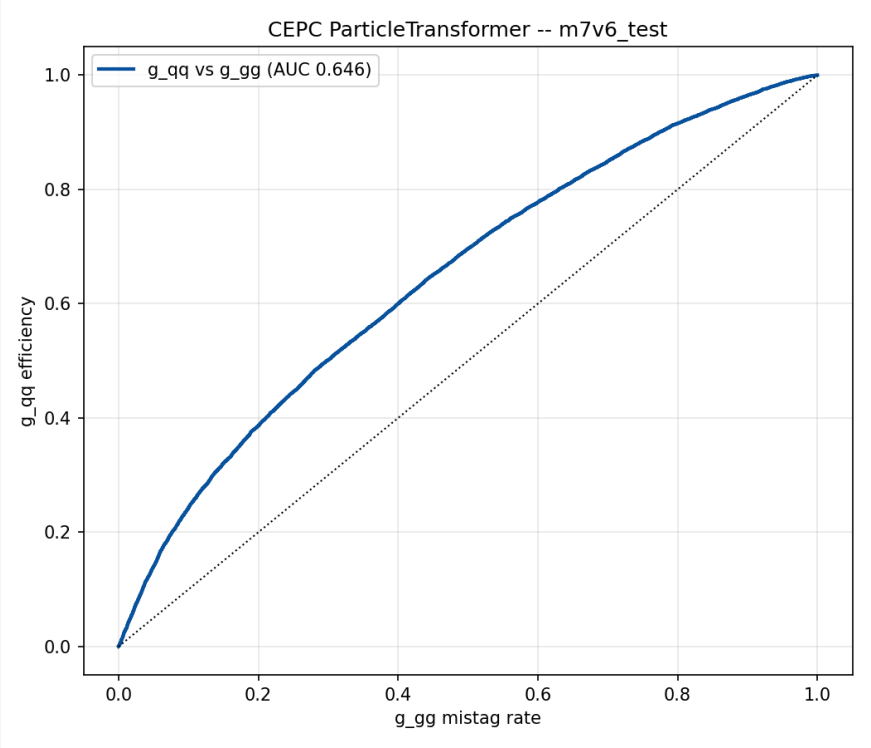


gggg	ggqq	qqqq
0.11	0.17	0.35

Dilution ⟨cos2Δφ⟩ between truth and reco. The origin of the offset and the wide migration is still under study.

Backup: tagger performance, S. Younes model





AUC g_qq vs g_gg

0.65

g_gg recall

0.47

g_qq recall

0.40

Held-out test set. Model trained before the label fix: g_qq is 26% of gluon jets, against about 11% with the corrected label.