Measurement of spin correlations induced by gluon polarization in parton showers with CMS

arXiv:2603.03689 CMS-SMP-25-006

Shuangyuan Wang on behalf of CMS collaboration
Zhejiang University

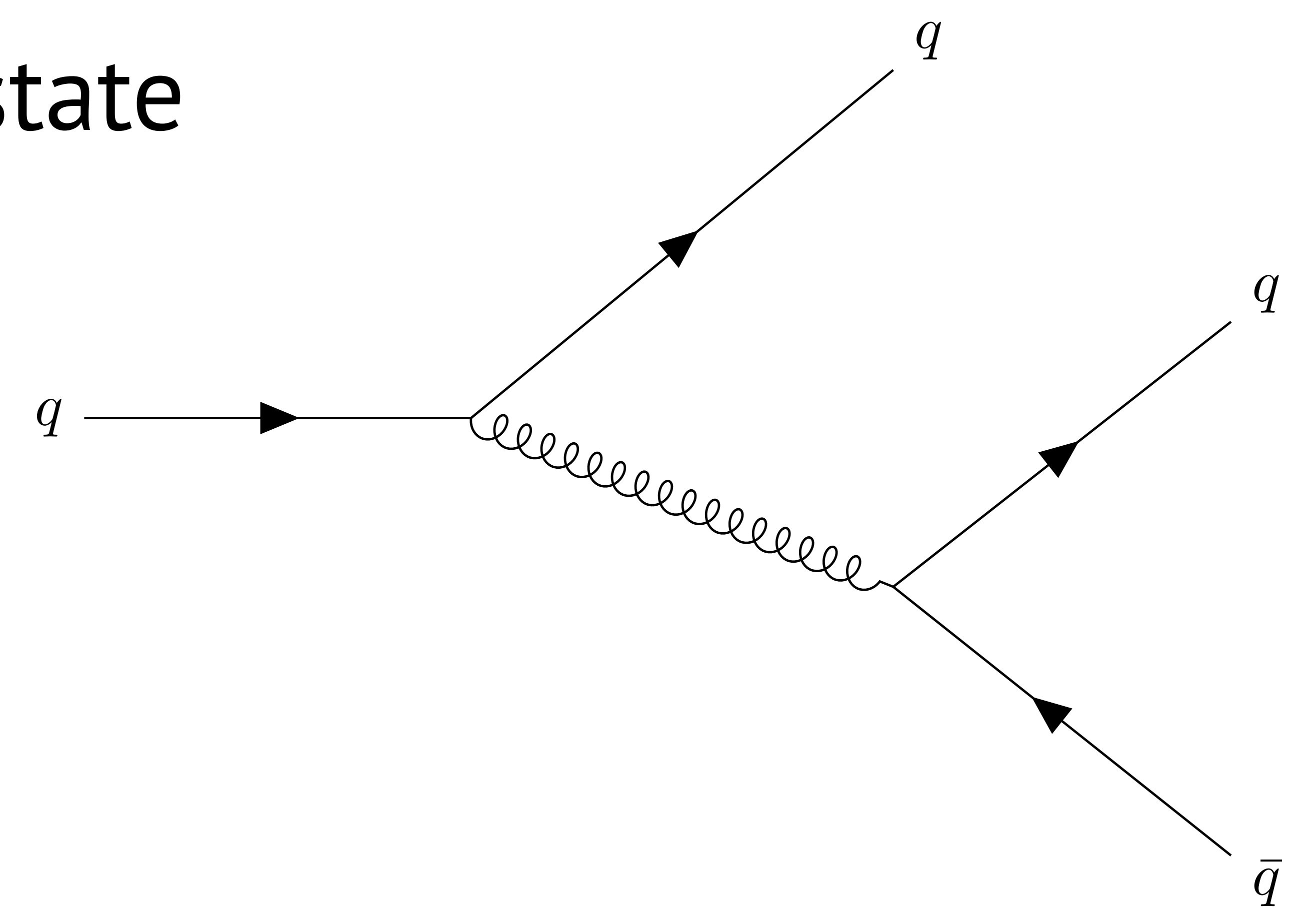
July 15, 2026

High Energy Physics Branch of CPS

Introduction

Gluon: spin-1 particle, quantum interference when produced in intermediate state

$$|M_+ + M_-|^2 = |M_+|^2 + |M_-|^2 + 2\,|M_+||M_-|\cos 2\Delta\phi$$



Effect at matrix element level observed at LEP, with jet-jet correlation

Effect in collinear and soft parton splitting?

- Never measured in data before

- Non-perturbative effect large, may change the effect

- Only possible to measure at the LHC: partons carry higher energy and split more significantly

Why measure spin correlation in PS

From Gregory Soyez
NLL PS developments

PanScales

- ▷ Parton showers Beyond leading logarithmic accuracy [2002.11114]
- ▷ PanScales parton showers for hadron collisions: formulation and fixed-order studies [2205.02237]
- ▷ PanScales showers for hadron collisions: all-order validation [2207.0946
- ▷ Next-to-leading logarithmic PanScales showers for Deep Inelastic scatte and Vector Boson Fusion [2305.08645]
- ▷ Colour and logarithmic accuracy in final-state parton showers [2011.10(
- ▷ Spin correlations in final-state parton showers and jet observables [2103
- ▷ Soft spin correlations in final-state parton showers [2111.01161]
- ▷ Introduction to the PanScales framework, version 0.1 [2312.13275]

Various combinations of M.vanBeekveld, M.Dasgupta, B.El Menoufi, S.Fe Ravasio, K.Hamilton, J.Helliwell, A.Karlberg, R.Medves, P.Monni, G.P.Sa L.Scyboz, A.Soto-Ontoso, G.Soyez, R.Verheyen

Deductor

- ▷ Summations of large logarithms by parton showers [2011.04773]
- ▷ Summations by parton showers of large logarithms in electron-positron annihilation [2011.04777]

Z.Nagy, D.Soper

Alaric

- ▷ A new approach to color-coherent parton evolution [2208.06057]
- ▷ New approach to QCD final-state evolution in processes with massive partons [2307.00728]
- ▷ alaric parton shower for hadron colliders [2404.14360]

Combinations of B.Assi, F.Herren, S.Höche, F.Krauss, D.Reichelt, M.Schönherr

Apollo

- ▷ A partitioned dipole-antenna shower with improved transverse recoil [24

C. Preuss

CVolver

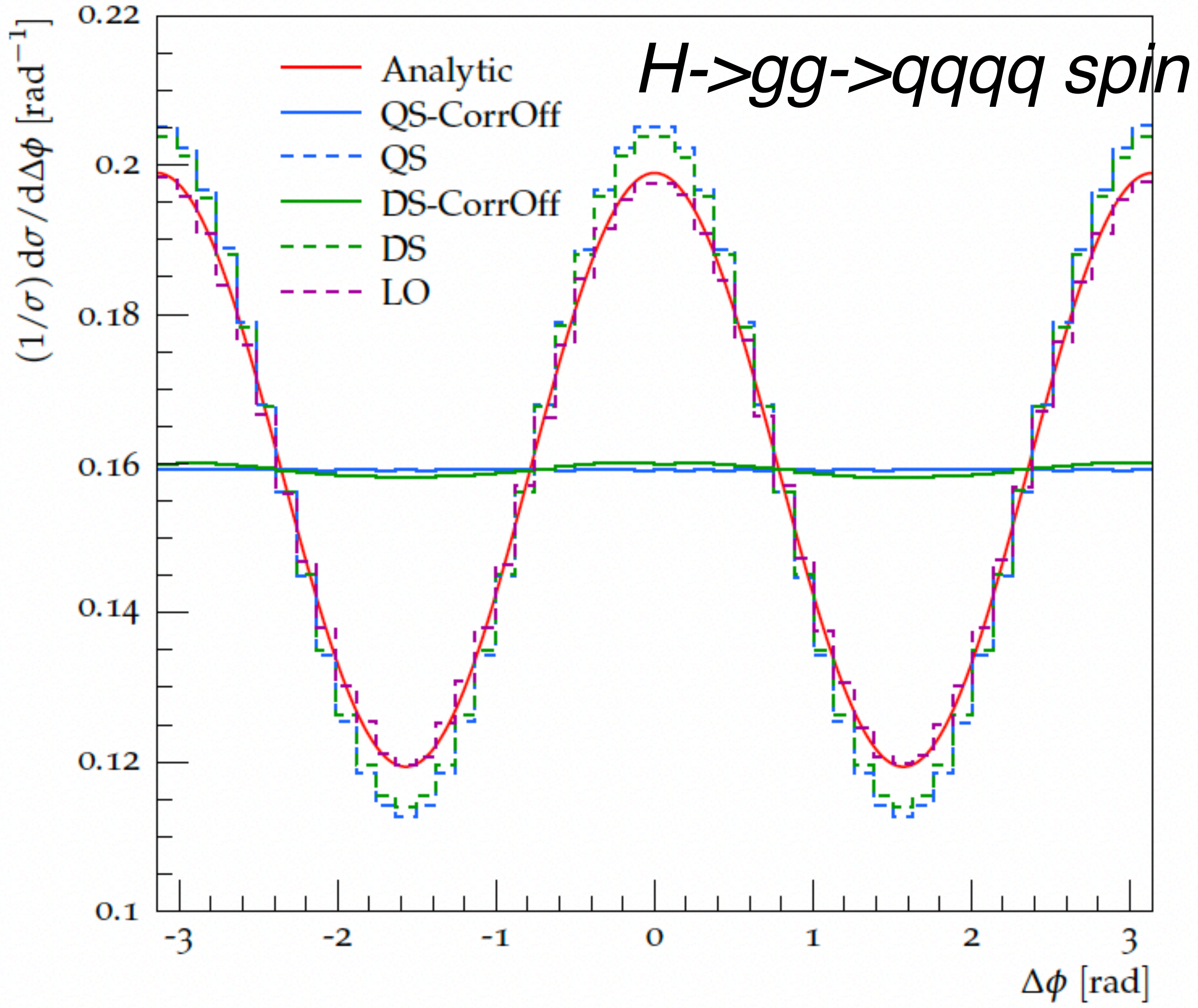
- ▷ Parton branching at amplitude level [1905.08686]
- ▷ Building a consistent parton shower [2003.06400]
- ▷ Improvements on dipole shower colour [2011.15087]

J.Forshaw, J.Holguin, S.Platzer

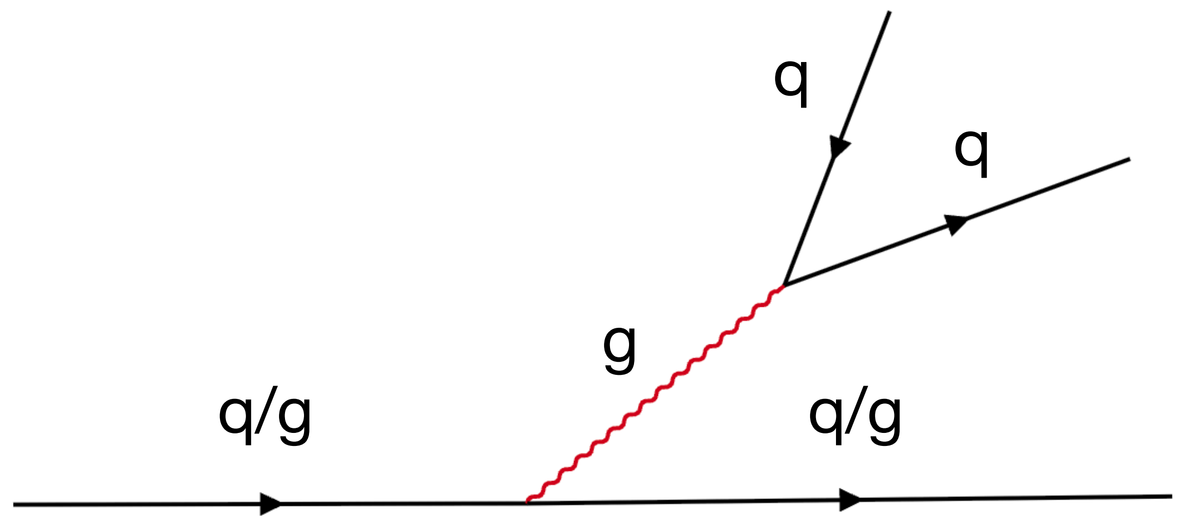
- **Check PS models**
  - Mainstream PS models at LL accuracy
    - Pythia/Sherpa: spin averaging, add a phi modulation
    - Herwig: explicit spin correlation tracing at collinear limit
  - Full spin correlation is a vital ingredient to NLL PS
    - NLL has become the standard in the PS community
    - Included in PanScales

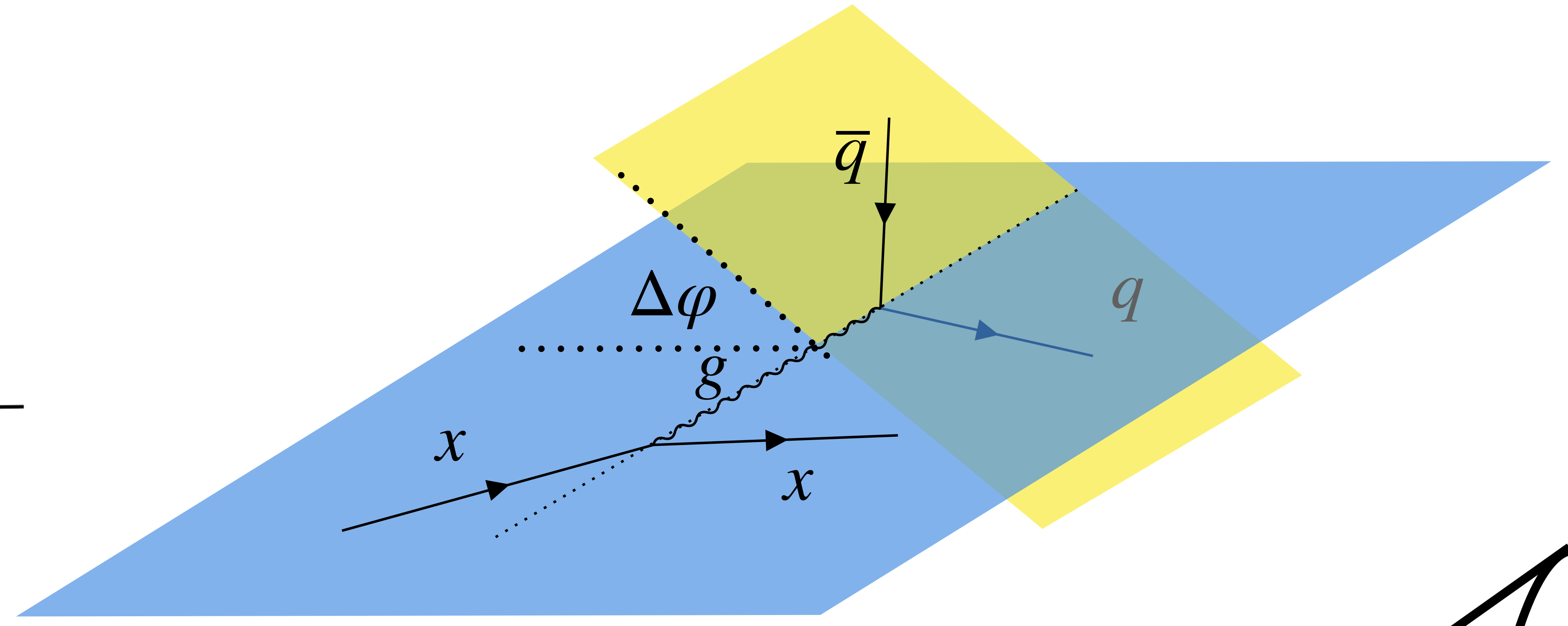
- **Wrong spin correlation biases machine learning in jet substructure**
  - QCD often important bkg for boosted object tagging
  - Large effect in H->gg splittings, will be interesting to apply this analysis method to boosted H->gg searches in the future

H->gg->qqqq spin correlation



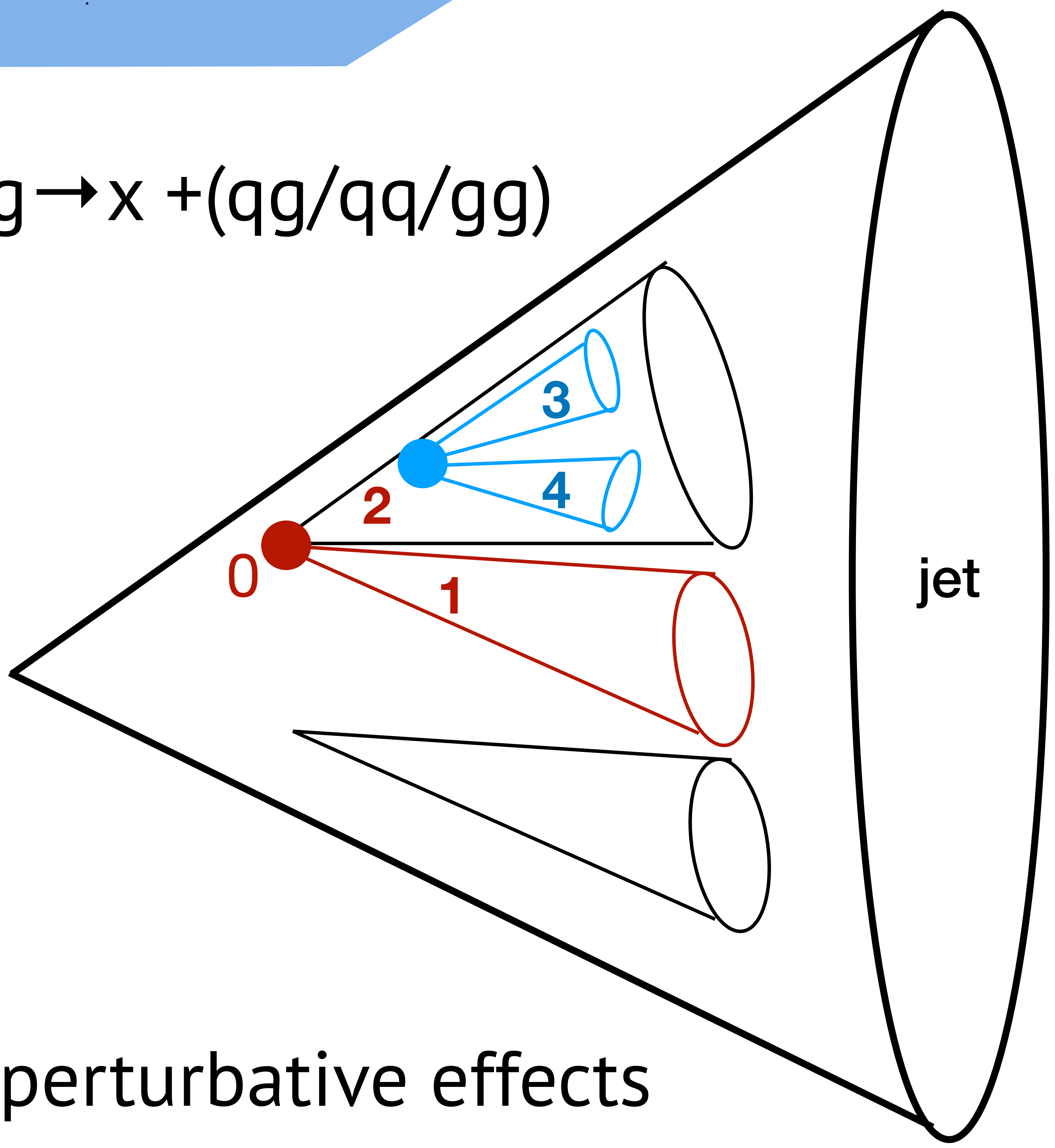
Analysis overview





Spin-sensitive observable $\Delta\varphi$ arXiv:1807.01955

- An intermediate particle produced and decay in PS: x→x+q/g→x +(qg/qq/gg)
- $\Delta\varphi$: Angle between production and decay plane in PS



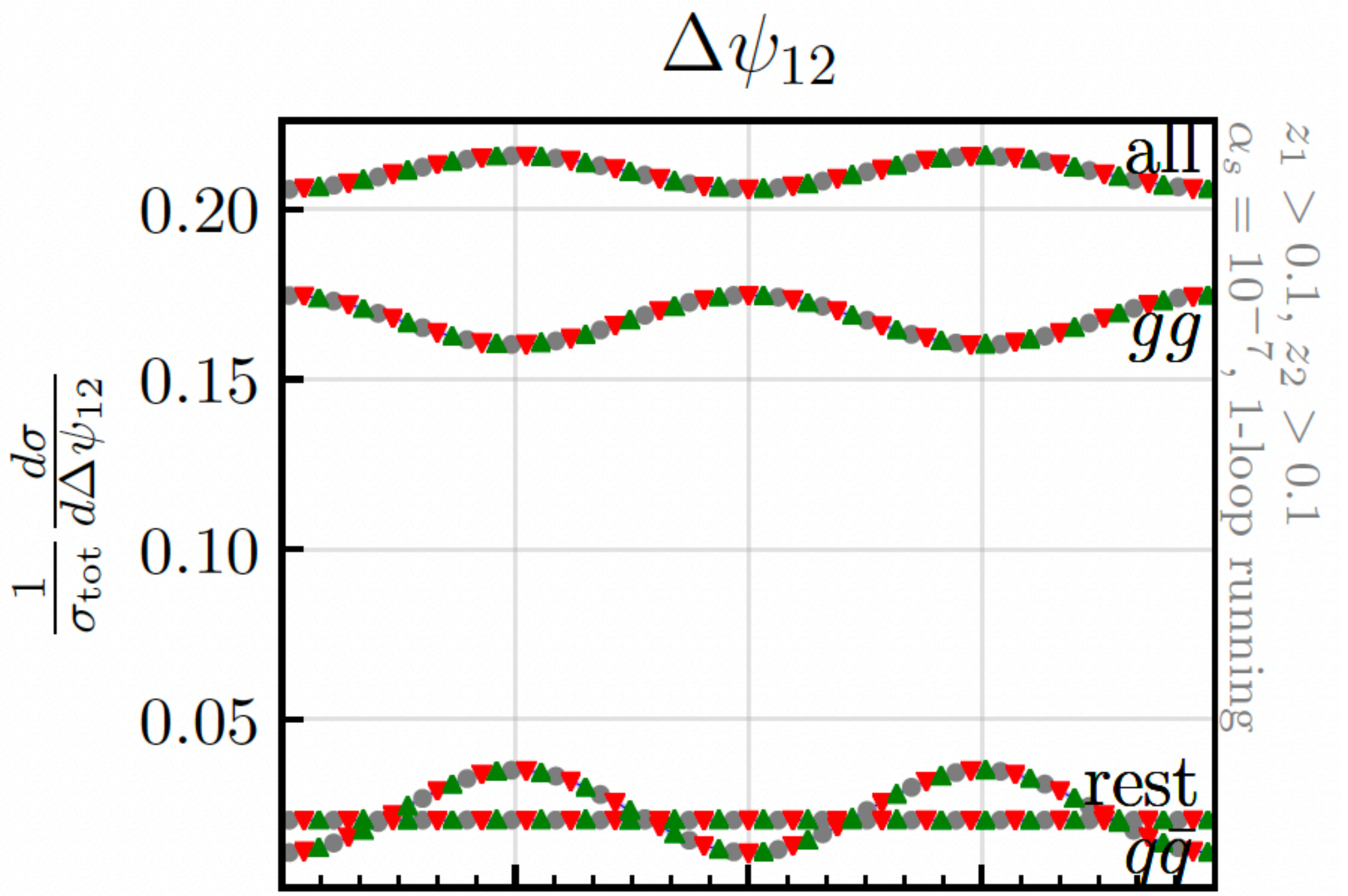
Analysis target:

- Find subjets representing parton splitting
- Measure $\Delta\varphi$ of an intermediate gluon in a jet using subjets
  - The $\Delta\varphi$ between plane of subjet 1,2 and subjet 3,4
- Check the size of spin correlation at particle level with non-perturbative effects
- Provide data for MC description

Analysis strategy

Theoretical prediction at parton level

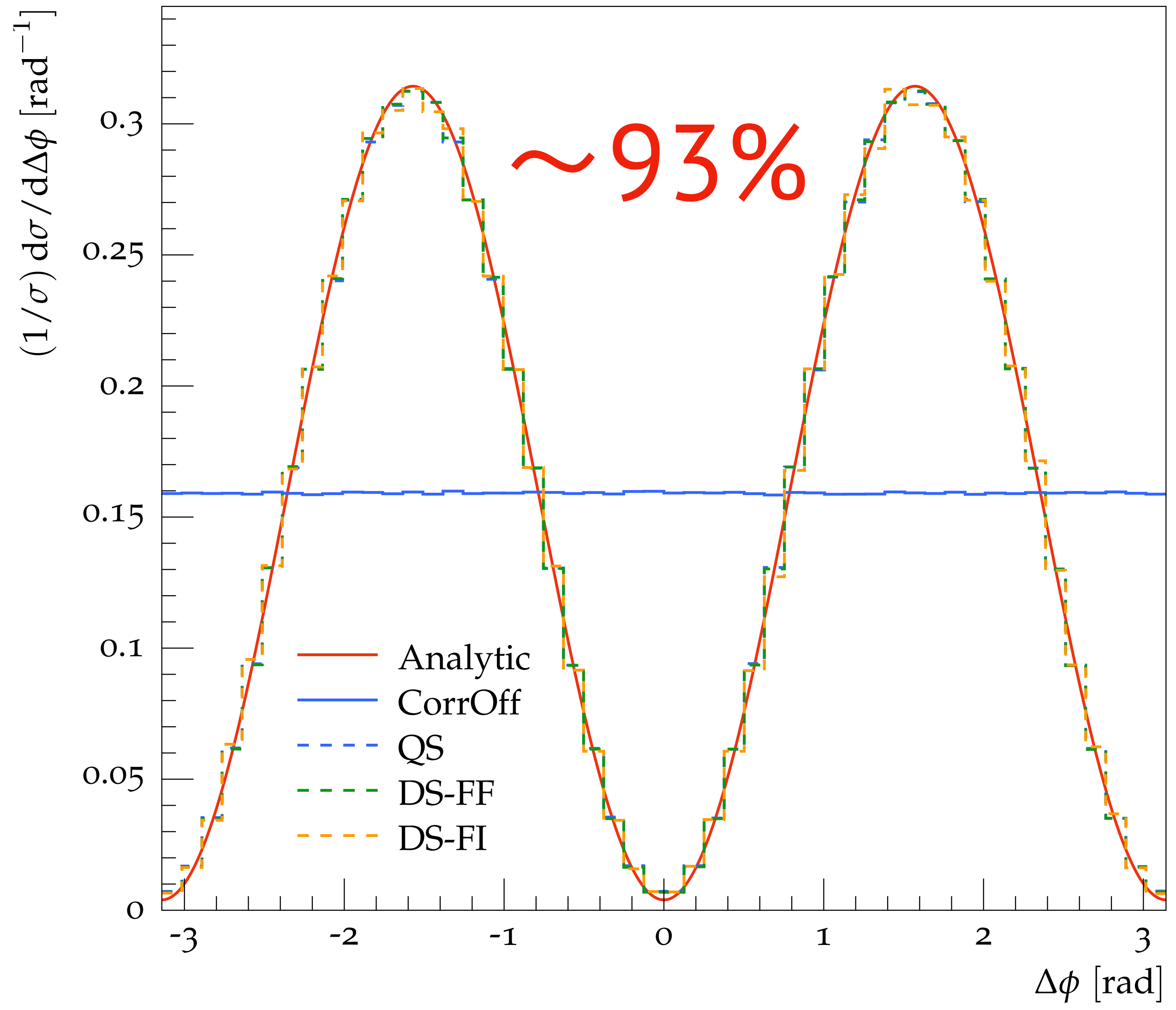
arXiv:2103.16526, in Panscale



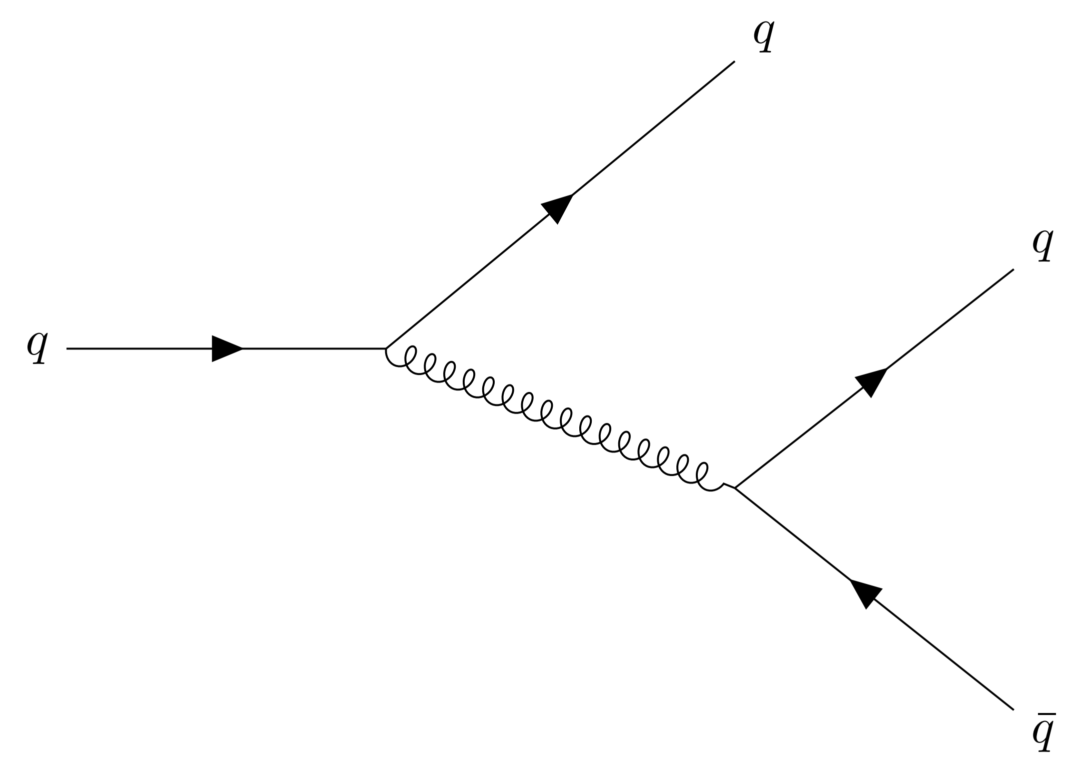
Cancellation of g->qq and g->gg small inclusive effect

arXiv:1807.01955v1, in Herwig

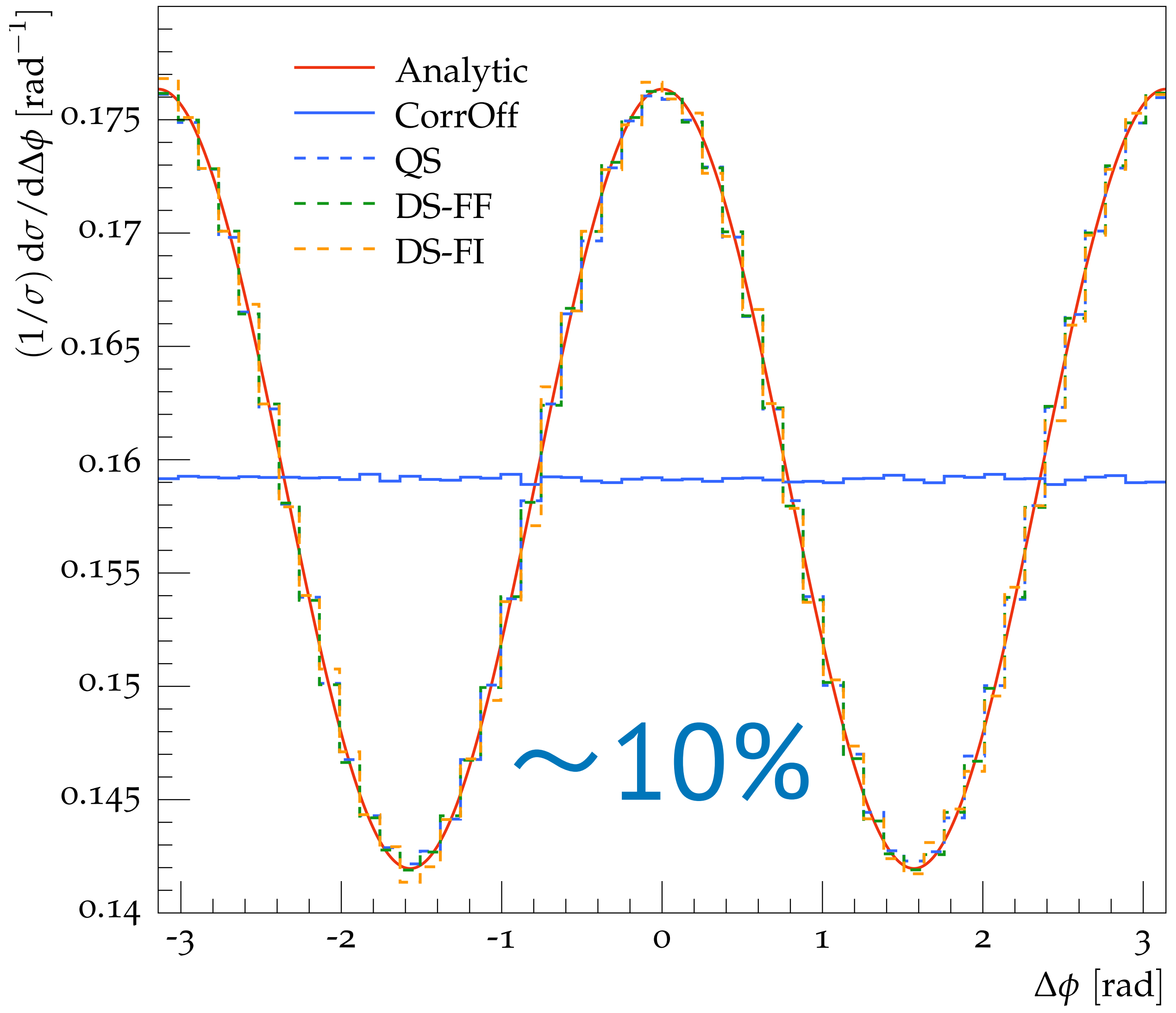
q⟶q+g⟶q+(q+q)



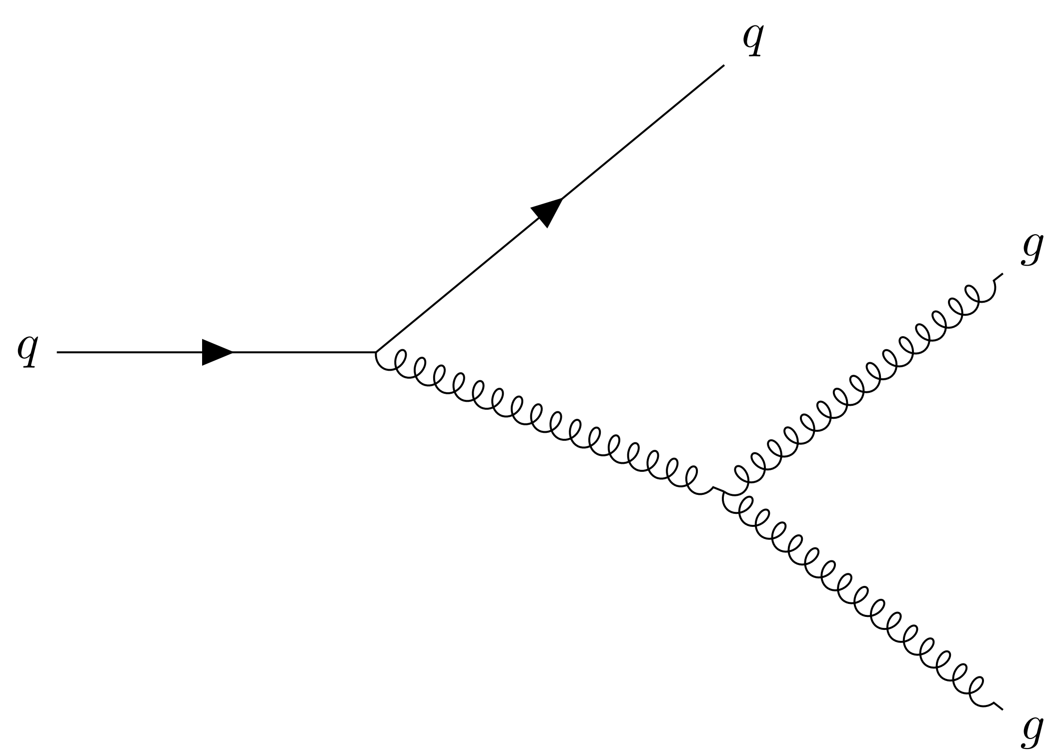
Important to tag this one



q⟶q+g⟶q+(g+g)

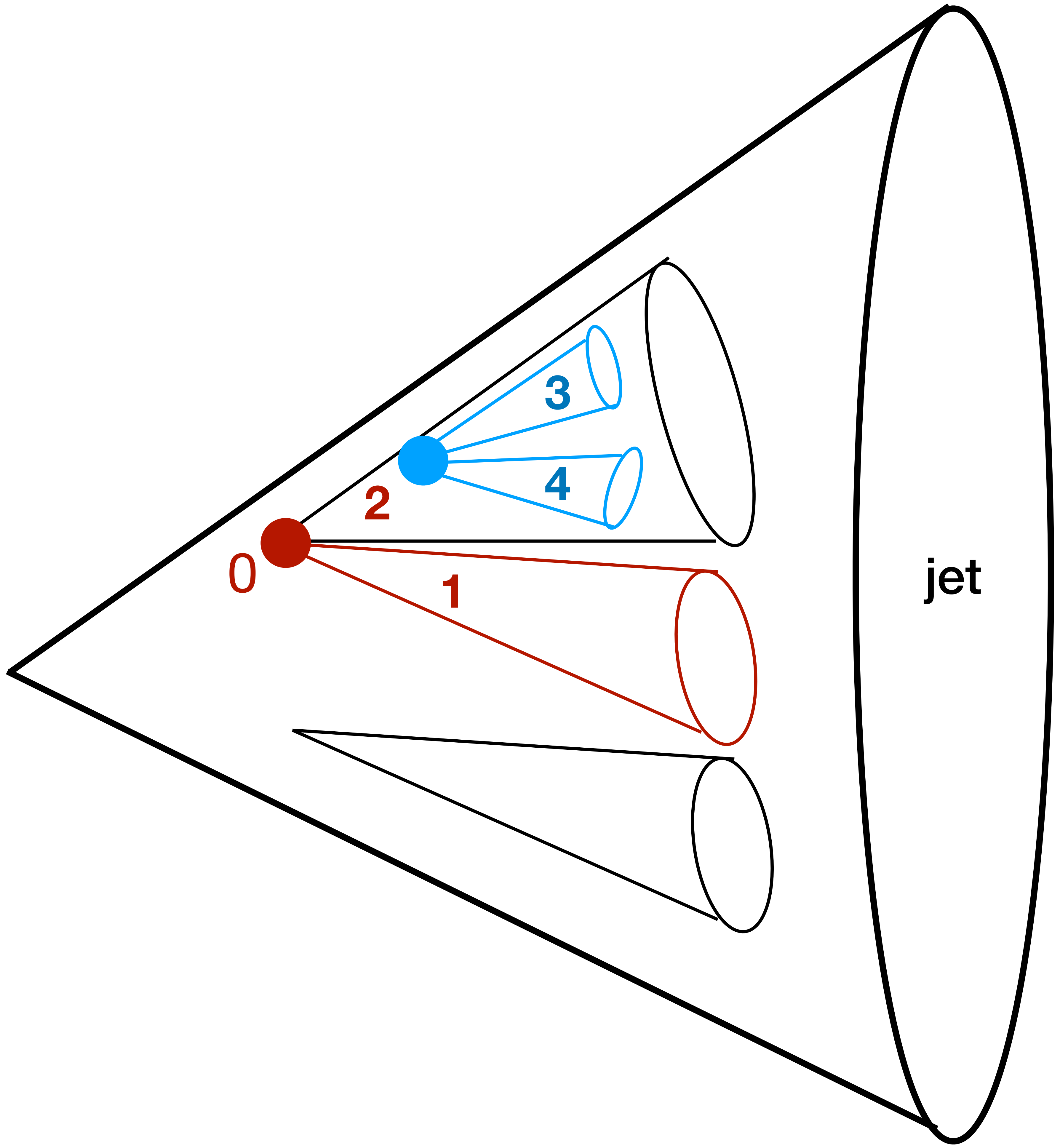


Opposite sign to g → qq



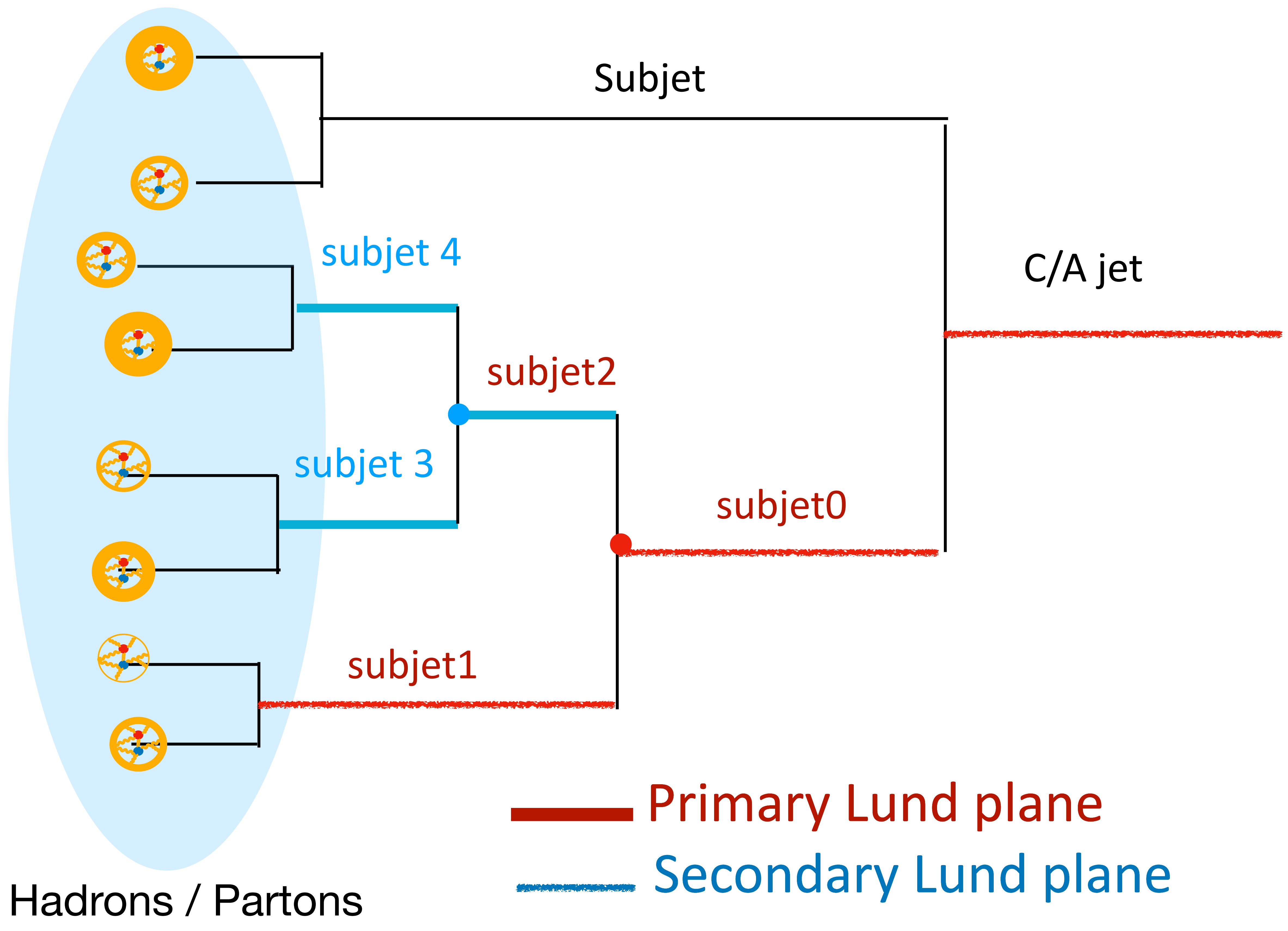
Dataset & selection

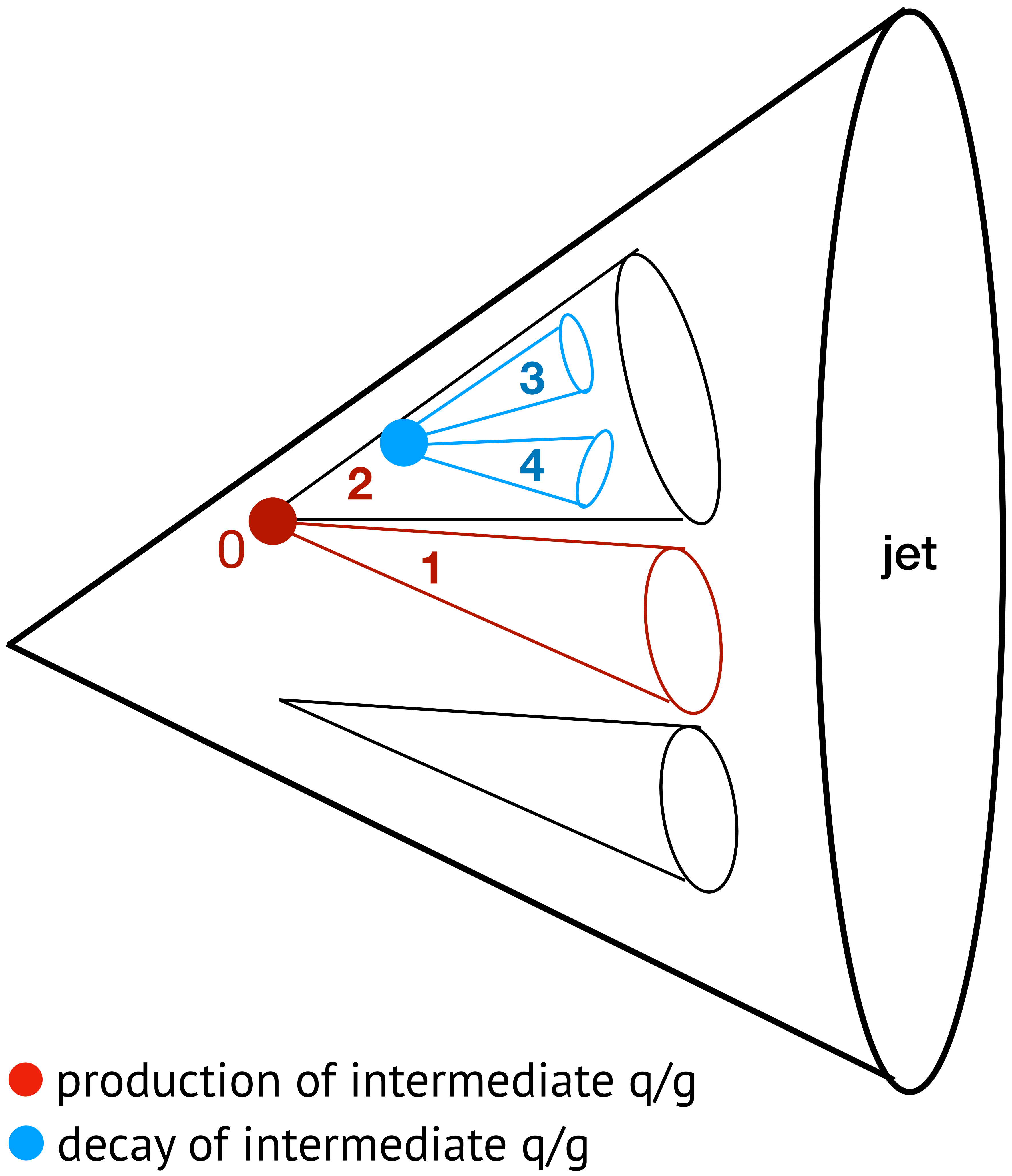
- pp data: 34.7 fb^{-1} in 2022 at $\sqrt{s}$ = 13.6 TeV
- AK8 jets: $\mathrm{p}_{\mathrm{T}}^{\mathrm{jet}}$ > 700 GeV, $|\eta_{\mathrm{jet}}|$ < 1.7
- C/A declustered subjet: all particles pT > 1 GeV
- Simulation:
  - PYTHIA 8.306, HERWIG 7.3.0
  - Correlation on/off



Analysis strategy

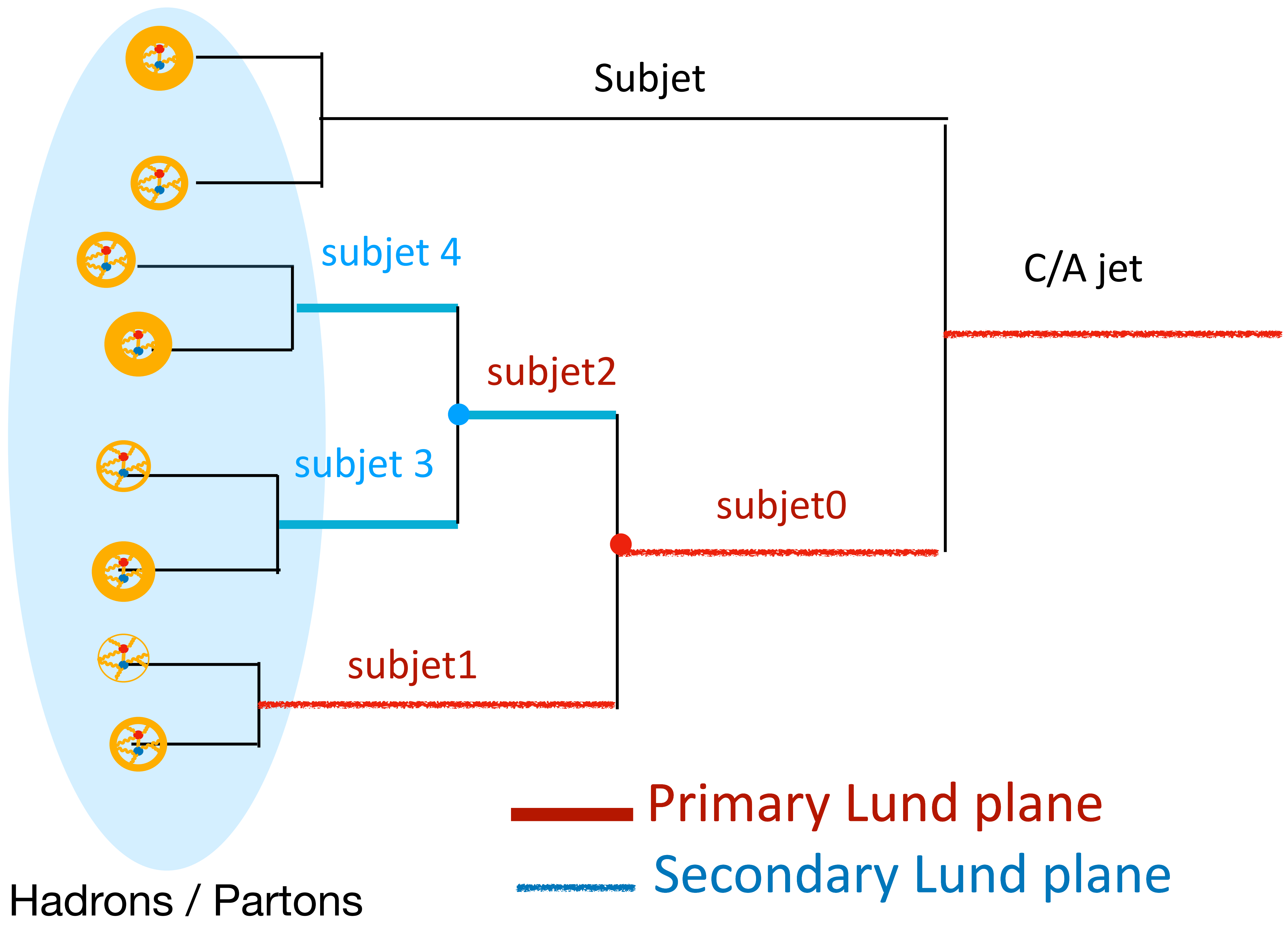
- Step 1: Restore parton splitting in jets
    - C/A decluster





Analysis strategy

- Step 2: In each jet, select subjets representing splitting of 0 -> 1, 2 (2->3,4)



- 0->1,2, max k_T in primary lund plane
- 2->3,4, max k_T in secondary lund plane

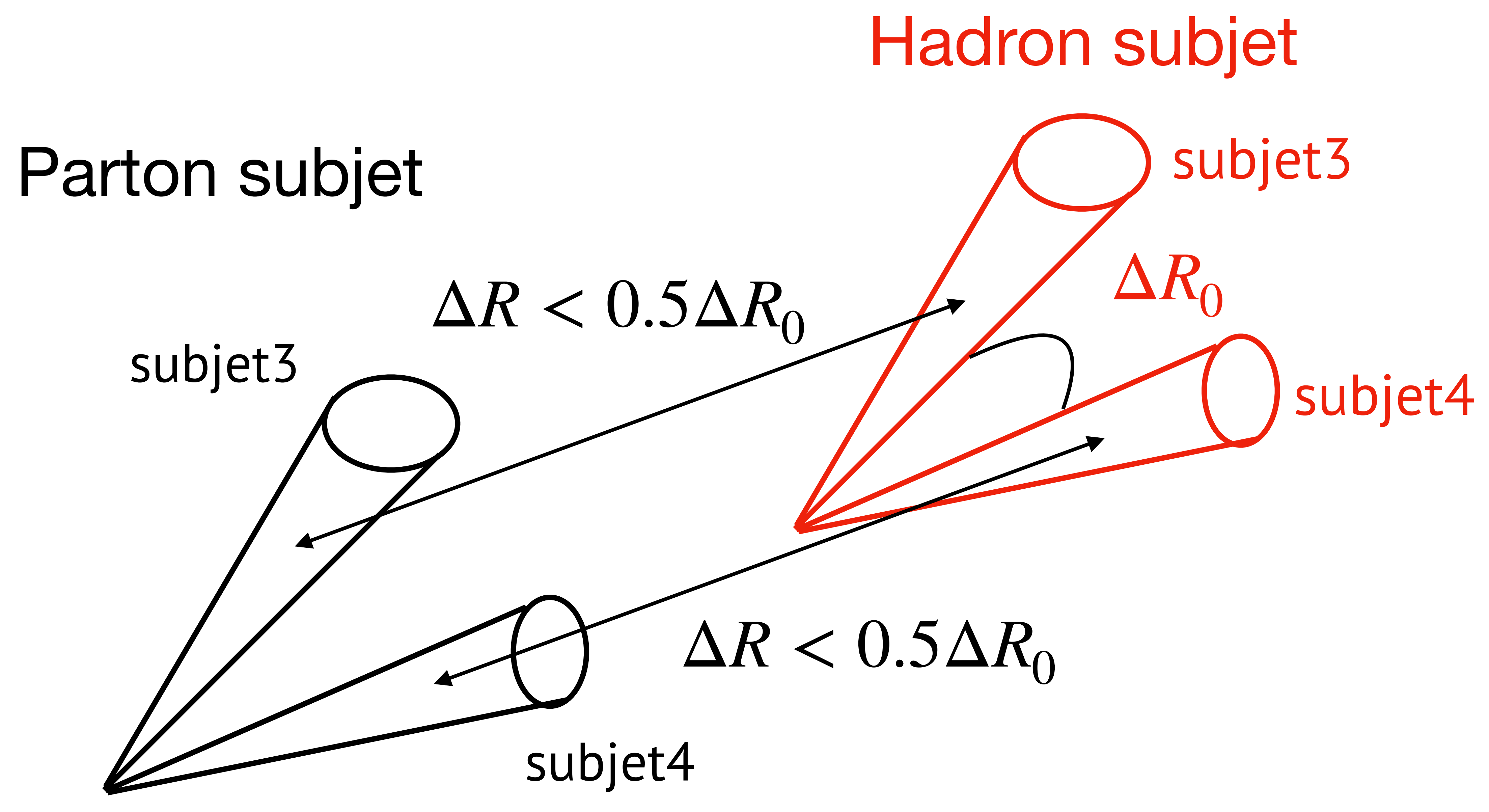
arXiv:2111.01161

- $P_T^{\mathrm{subjet2}} > 130$ GeV, enough particles for decluster and flavor tagging
- Selected splittings 0->1,2(2->3,4) could be:
  - $\mathrm{x} \to \mathrm{xq}(\mathrm{q} \to \mathrm{qg})$
  - $\mathrm{x} \to \mathrm{xg}(\mathrm{g} \to \mathrm{gg})$
  - $\mathrm{x} \to \mathrm{xg}(\mathrm{g} \to \mathrm{q\bar{q}})$
- need flavor tagging on subjet 3&4

Analysis strategy

- Step 3: Flavor tagging to distinguish different processes
  - Machine learning, DNN

- Hadron subjets true flavor: matching to parton subjets
  - same C/A decluster and selection on partons
  - match parton and hadron subjet 3&4
  - not guaranteed to find one
- Parton subjets flavor
  - IFN algorithm: infrared and collinear safe definition



arXiv:2306.07314

Analysis strategy

- Step 3: Flavor tagging to distinguish different processes
    - Machine learning, DNN

- Four classes
    - unmatched
    - $g \to gg$
    - $g \to q\bar{q}$
    - $q \to qg$

- Training input, HERWIG correlation off simulation
    - Splitting features

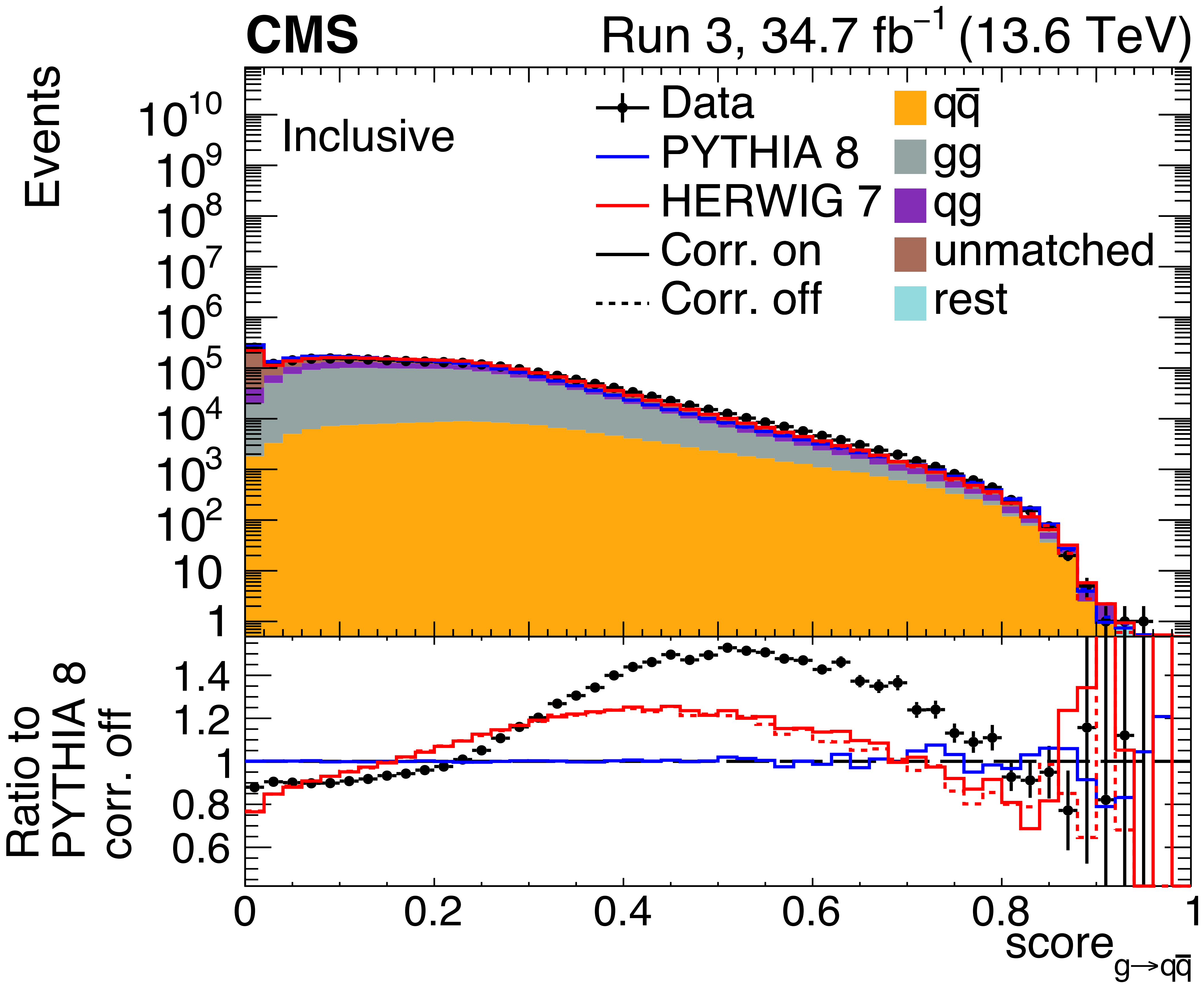
      z_1: $p_{T_2}/(p_{T_1}+p_{T_2})$, z_2: $p_{T_4}/(p_{T_3}+p_{T_4})$

      k_{T2}: k_T of 2→3, 4 ΔR_{34}
    - Subjet 3&4 substructure

      #particles, #tracks

      σ: subjet width $p_T D$

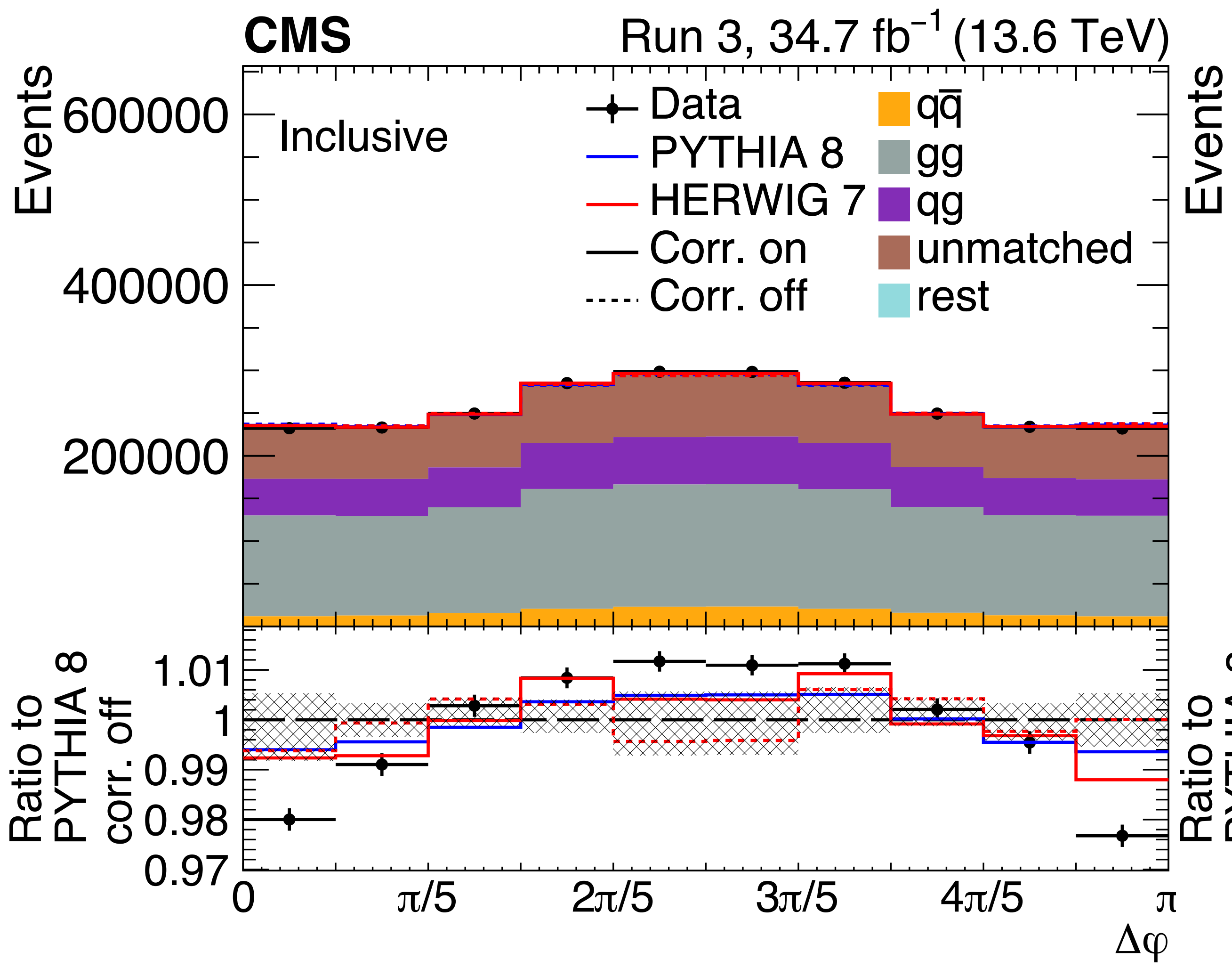
Output score



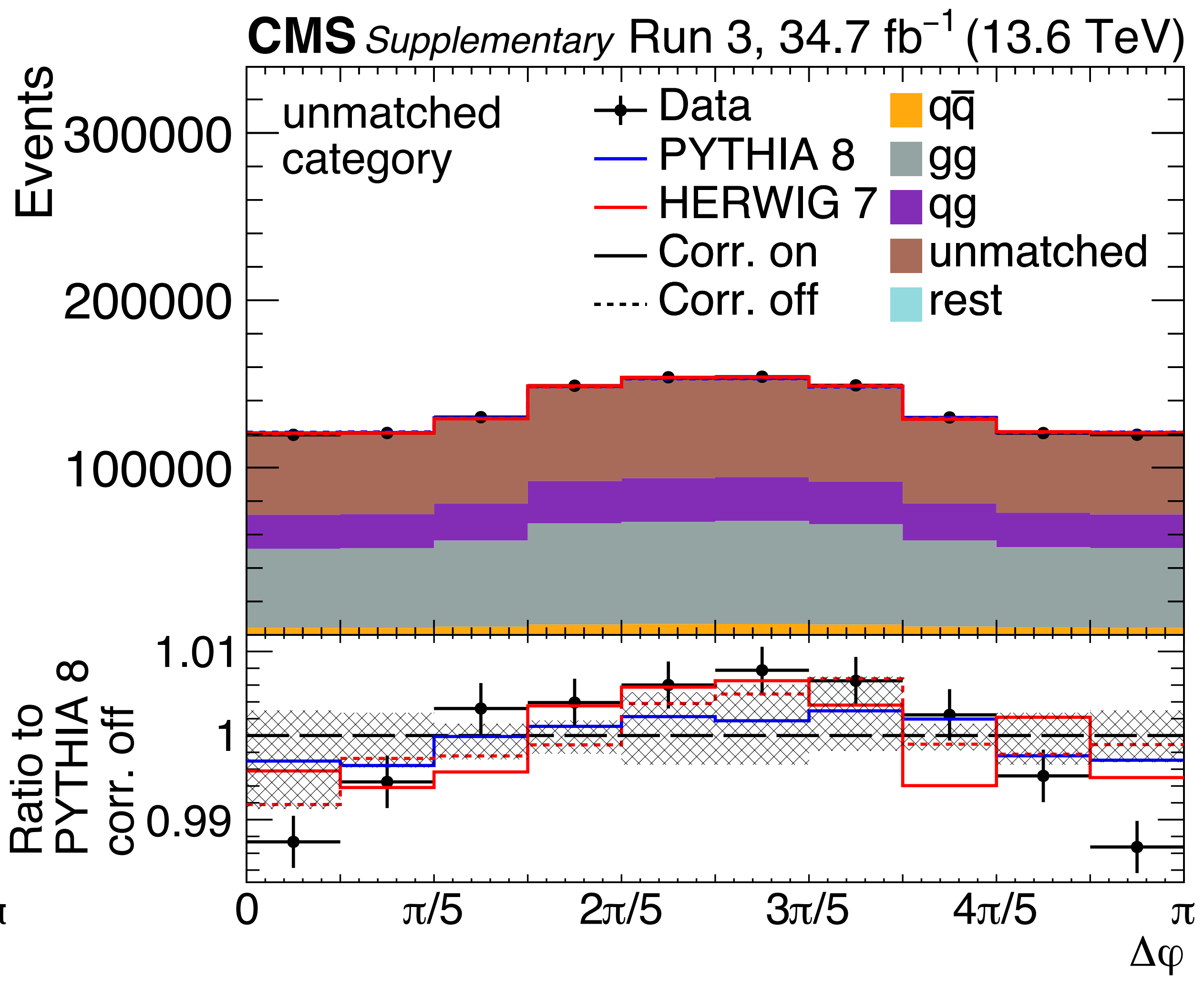
$\mathrm{score}_{g\to q\bar{q}}$

- Score distinguishes different processes
- Slight mismodeling in MC
  - MC-to-data score reweighting
- Events categoried to four classes according to max score
  - unmatched category
  - $q \to qg$ category
  - $g \to gg$ category
  - $g \to q\bar{q}$ category: used to measure spin correlation

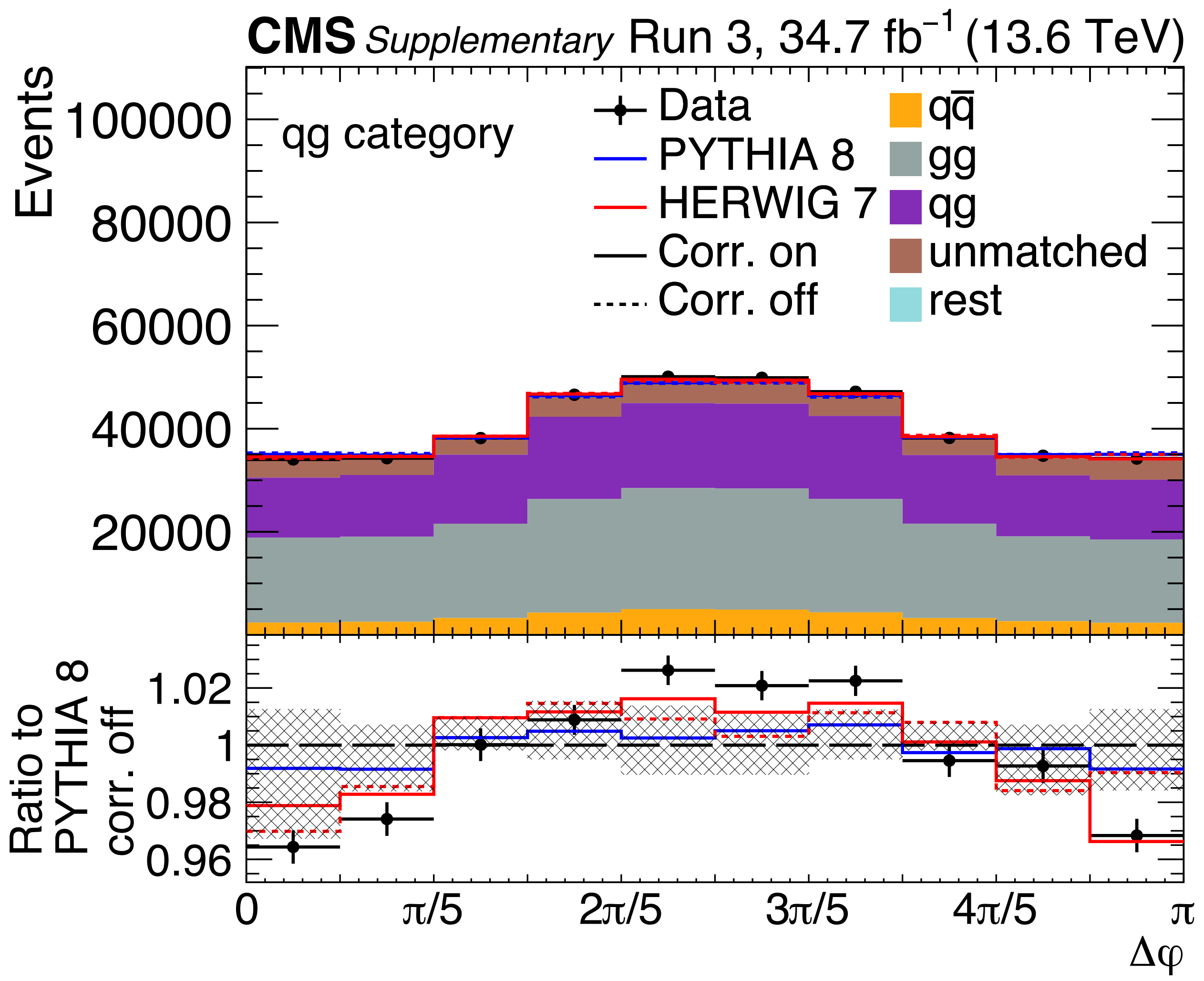
Spin effects in each category



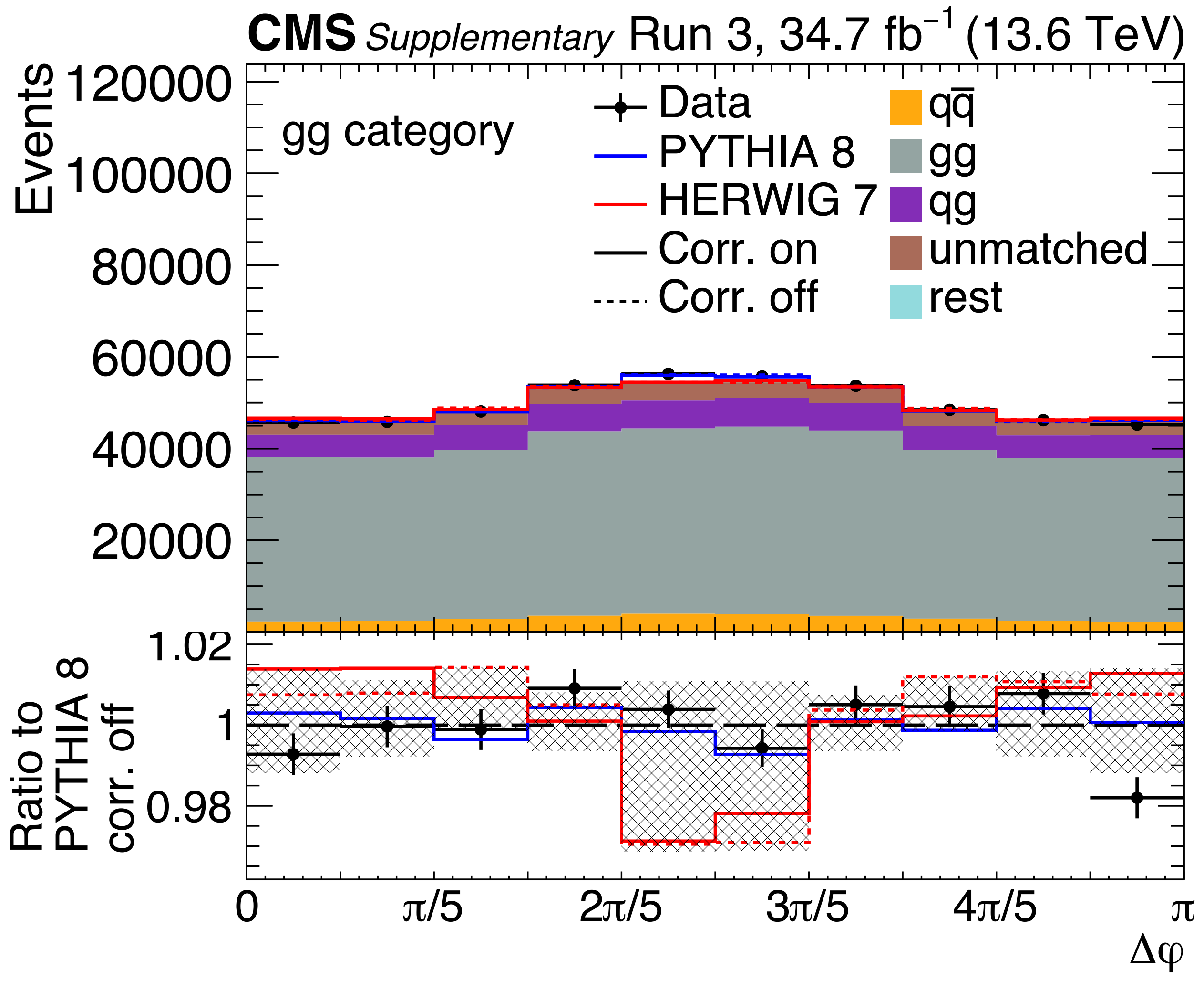
Inclusive



unmatched category



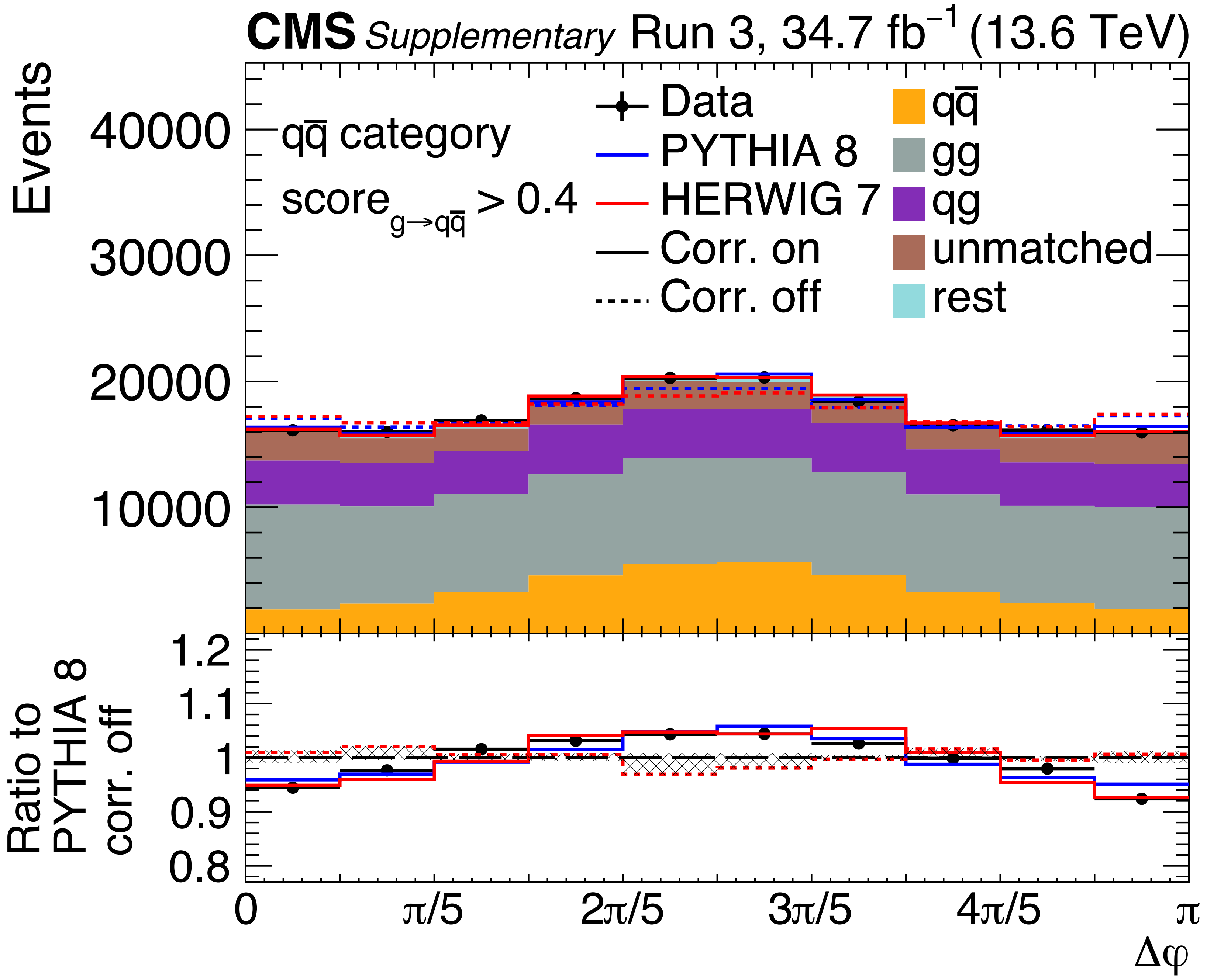
$q \to qg$ category

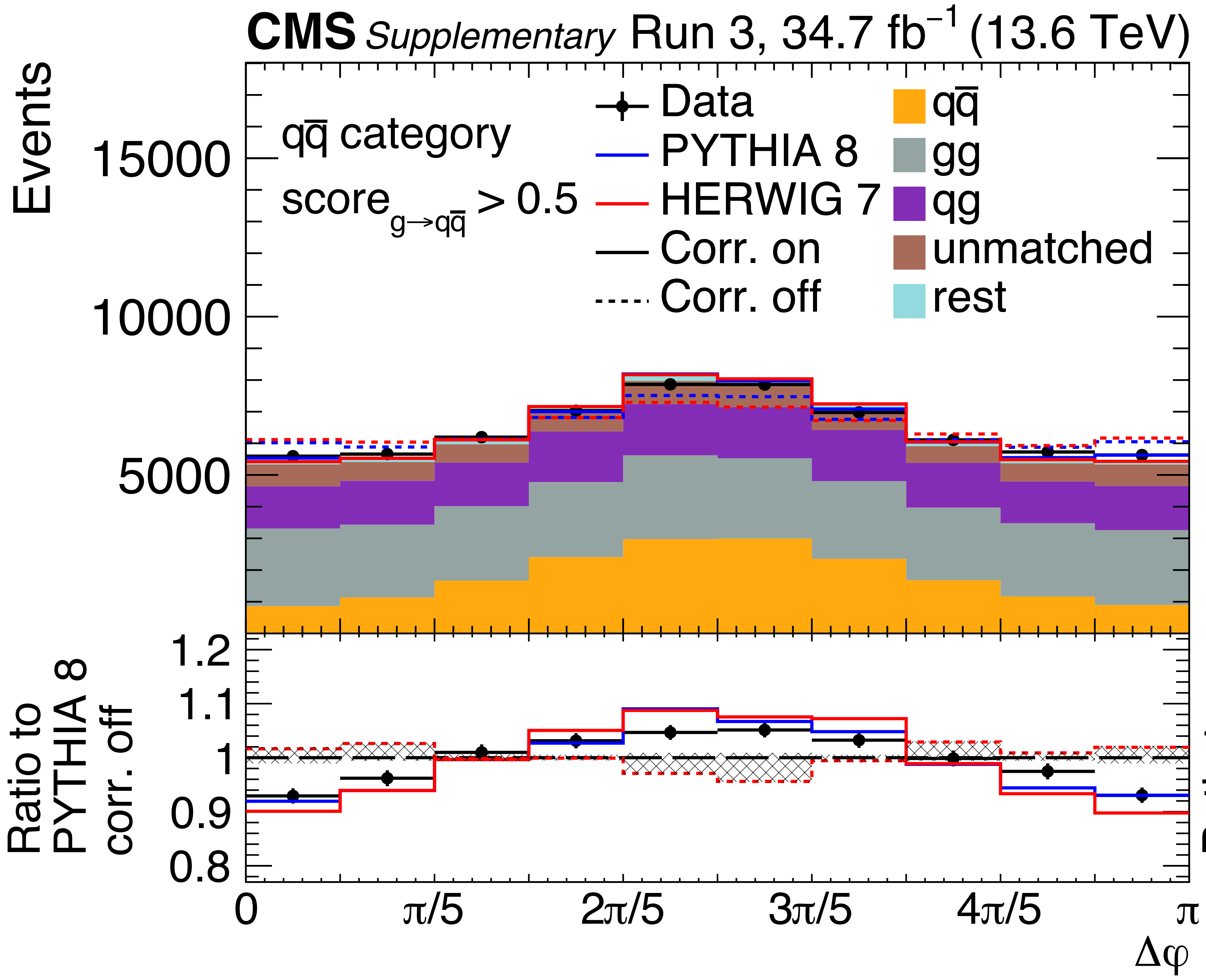


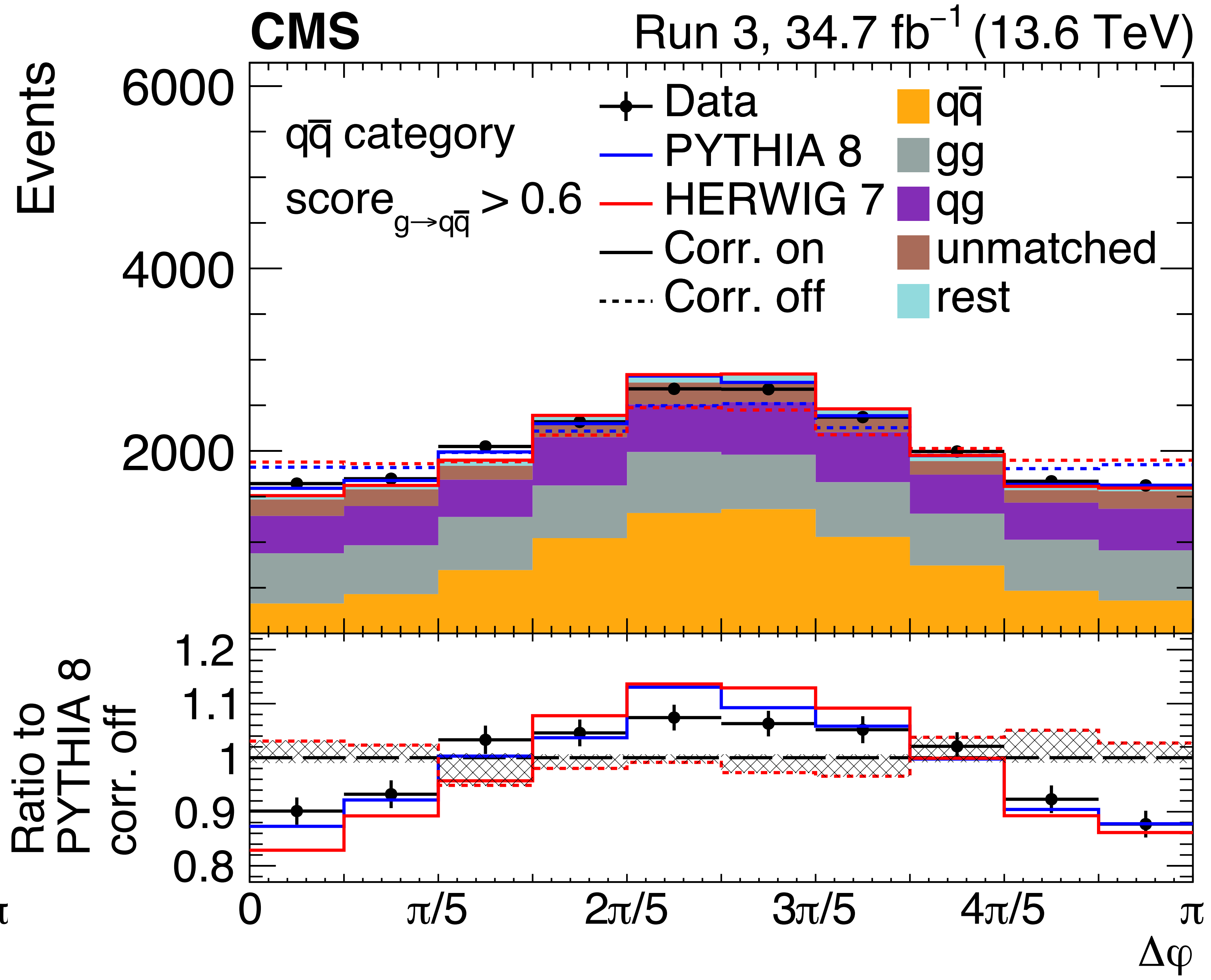
$g \to gg$ category

- Inclusive spin effect is 0.5% level, also tiny in non-qq categories
- Difference between Pythia and Herwig is small, most dominant in g->gg
- Data agrees better with Herwig in unmatched and qg, and Pythia in gg category

Spin effects in qq category







qq category

$\mathrm{score}_{g\to q\bar{q}} > 0.4$

qq category

$\mathrm{score}_{g\to q\bar{q}} > 0.5$

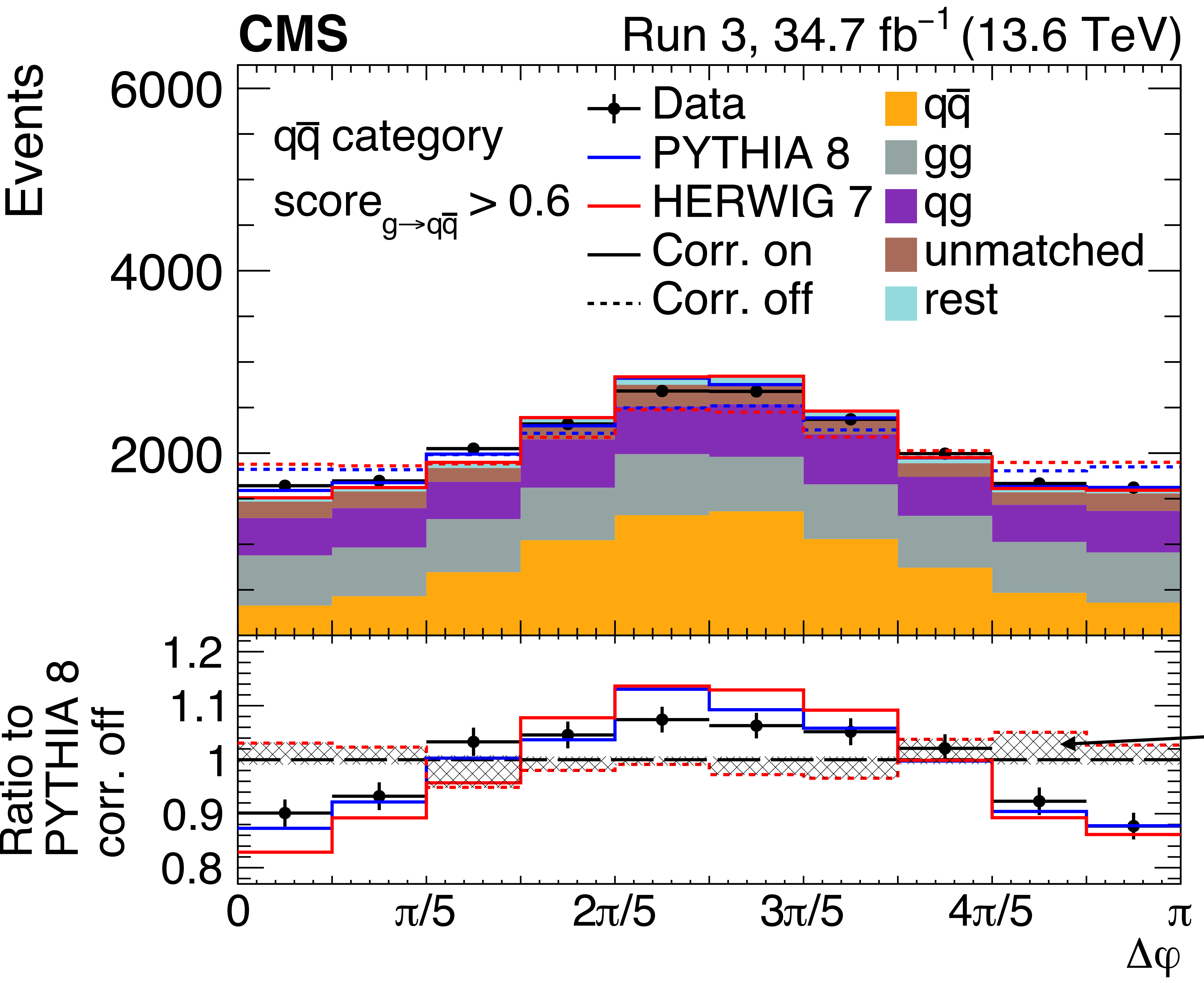
qq category

$\mathrm{score}_{g\to q\bar{q}} > 0.6$

With increasing score cut

- The purity of true qq events gets increased
- Spin effect enhanced in both Pythia and Herwig

Spin effects in qq category



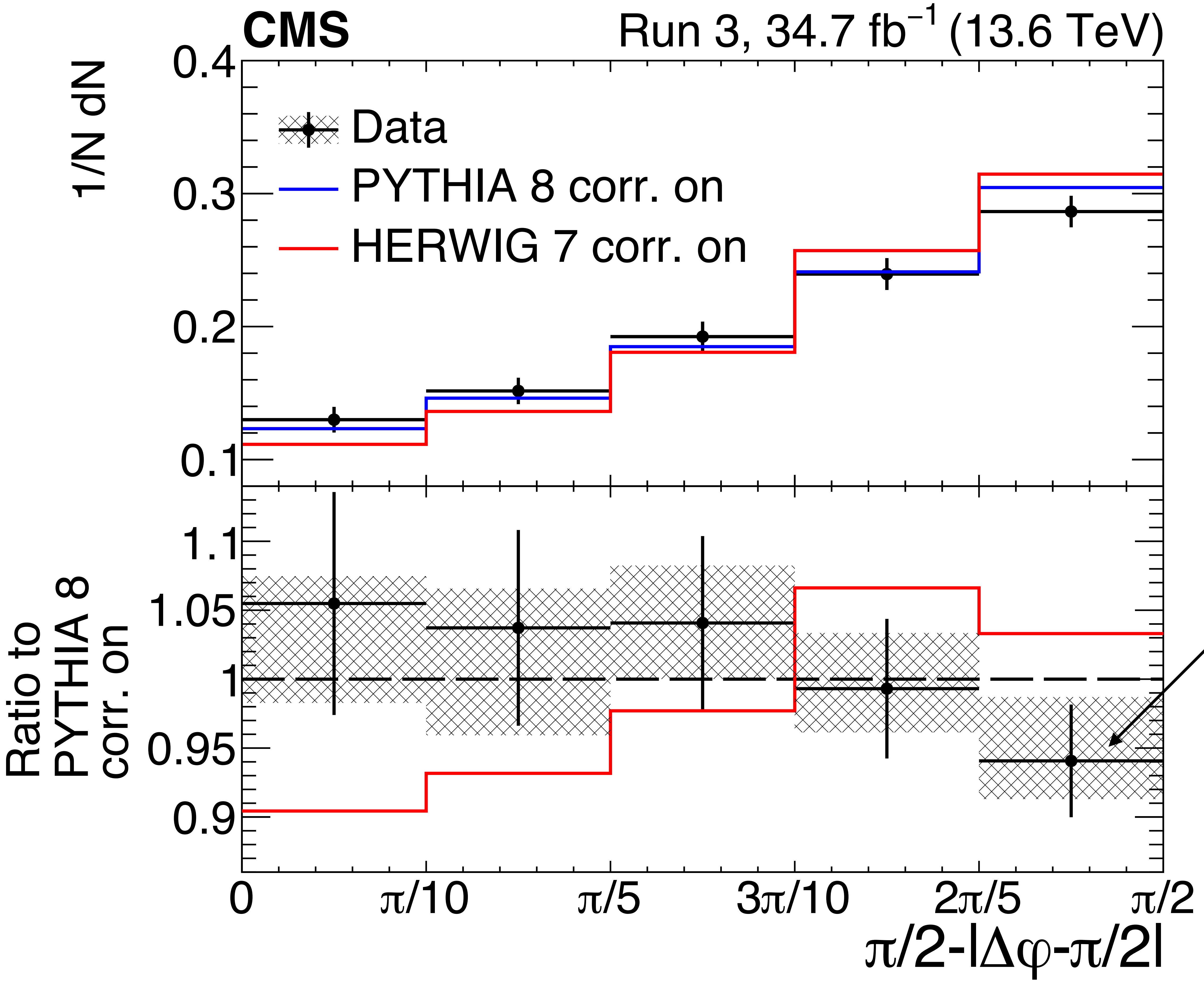
qq category

score$_{g\to q\bar{q}}$ > 0.6

- Final selection: score$_{g\to q\bar{q}}$ > 0.6
  - Enhance the purity of the qq class to ~ 38%
  - Enhance spin correlation effect

    Pythia ~ 12%, Herwig ~ 15%
- Data favors the correlation-on model, but still visible difference
- Leading uncertainties
  - FSR
  - MC stat
  - Pileup
- 6.8 σ from correlation-off, 1.9 σ from correlation-on

Unfolding



- Reco (gen) phase space:
  - score (g->qq) >0.6
  - jet pT > 700 (600) GeV
  - subjet2 pT> 130 (120) GeV
  - all particles pT >1 GeV
- Since $\Delta\varphi$ should be symmetric around π/2, we unfold $\pi/2 - |\Delta\varphi - \pi/2|$
- Leading uncertainties
  - modeling uncertainty (1-7%)

    Different unfolding results from PYTHIA and HERWIG.
  - tracking efficiency (1−3%)
- Data agrees better with PYTHIA

Summary

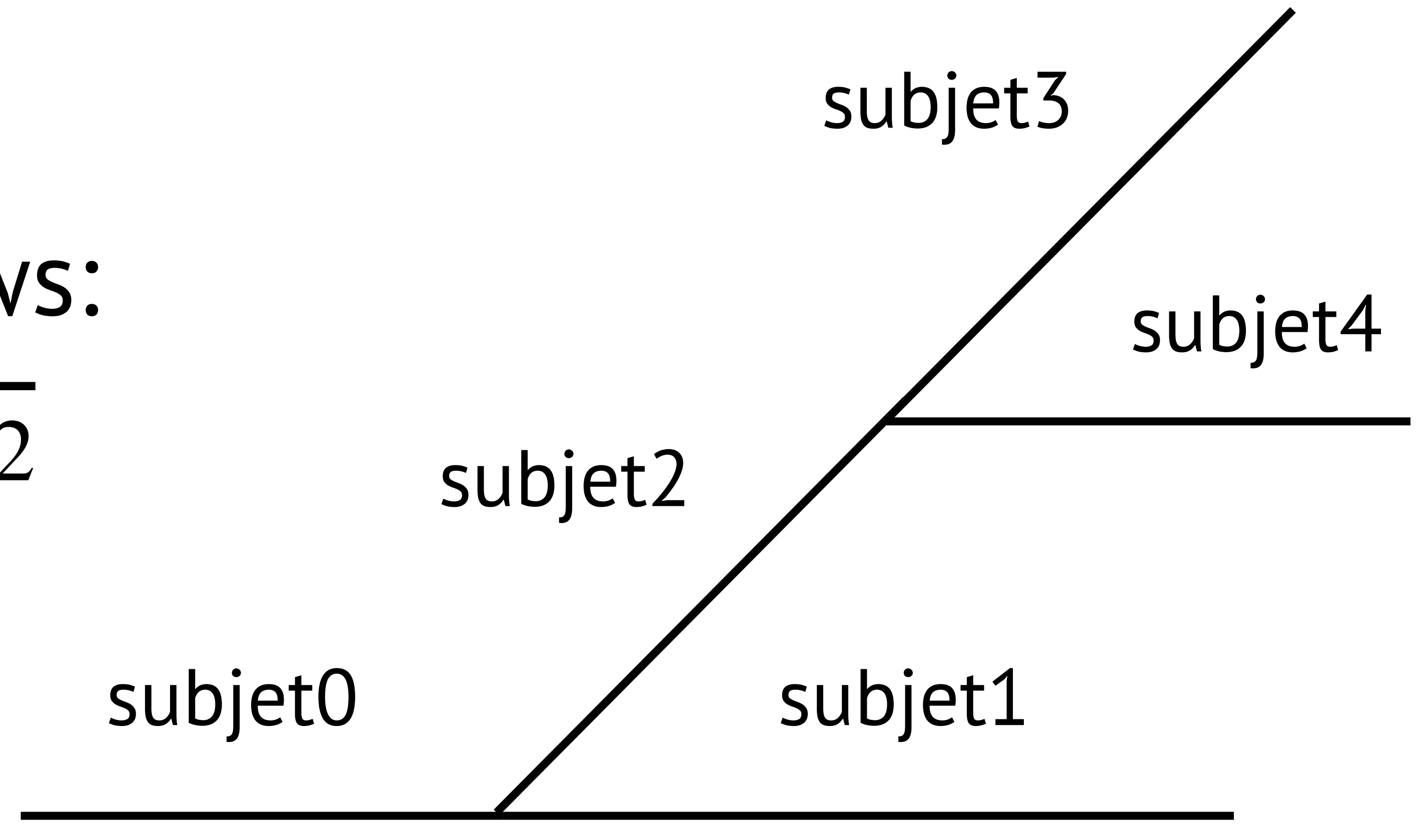
- First measurement of spin correlation in parton splitting inside jets is performed using CMS 2022 data
- 6.8σ significance against no spin correlation
- Visible spin correlation with non-perturbative effects
  - data indeed could be better modeled by taking into account the spin effects
- First attempt of light flavor tagging of parton splitting chains inside jet
  - Flavor dependent study of jet substructure can be explored in the future
- Interesting physics application: H->gg searches using correlations between gluon splittings
- HEPData Link

Backup

Input features

- $z_1 = p_{T_2}/(p_{T_1} + p_{T_2})$: The momentum fraction of the gluon in the first branching of $0 \to 1, 2$
- $z_2 = p_{T_4}/(p_{T_3} + p_{T_4})$: The momentum fraction of the soft subjet in the second branching of $2 \to 3, 4$
- ΔR: The angular distance between subjet 3 and 4 in $\eta - \phi$ plane
- k_{T2}: The k_T of secondary branching $2 \to 3, 4$, where k_T is defined as follows:

$$k_T = p_T^{softer} \cdot \Delta, \quad \Delta = \sqrt{(y_i - y_j)^2 + (\phi_i - \phi_j)^2}$$



- nparticles: Number of particles in a subjet
- ntracks: Number of charged particles in a subjet
- $p_T D$: Represents the distribution of p_T among particles within a jet, with form: $p_T D = \frac{\sqrt{\sum_i p_{Ti}^2}}{\sum_i p_{Ti}}$

σ: p_T weighted width of the subjet 3 and 4, defined as $\frac{\sum_i p_T^i \Delta R^i}{\sum_i p_T^i}$, where p_T^i is the p_T of particle i inside a subjet, and ΔR^i the angular distance between particle i and the subjet axis in $\eta - \phi$ plane.

Machine learning parameters

- Deep Neural Network

  - Hidden units: [256, 128, 128]
  - Dropout rate: 0.1

  - Learning rate: 0.0005
  - Batch size: 256
  - Epoch: 50

  - Loss function: categorical_crossentropy
  - Optimizer: Adam optimizer
  - Metrics: accuracy

  - Earlystopping monitor: val_loss
  - Earlystopping patience: 5
  - restore_best_weights: True

Uncertainties

- Theoretical Uncertainties:
  - Model uncertainty (~ 3%)

    difference between herwig and pythia
  - QCD scale of parton shower (~ 1%)

    μ_R is varied by a factor of 1/2 and 2 from nominal value for both ISR and FSR
  - QCD scale of hard scattering process (< 0.1%)

    μ_R and μ_F are varied by a factor of 1/2 and 2 from their nominal values.
  - PDF uncertainty (< 0.1%)

    due to the choice of PDF sets and value

- Experimental Uncertainties
  - MC stat uncertainty (~ 1.5%)
  - Pileup uncertainty (~ 1.1%)
  - DNN modeling uncertainty (~ 0.7%)
  - Energy Scale uncertainty of jet constituents:

    5% on neutral particles (~ 1.3%)

    3% on photons (~ 0.5%)

    1% on charged particles (~ 0.4%)
  - Track Efficiency (~ 0.5%)

    randomly lose 3% of tracks for jets and particles energy smearing by 10%
  - JES/JER uncertainty (~ 0.3%)