HACETTEPE UNIVERSITY

Monte Carlo Event Generator Tuning at CMS

Mehmet Umut Ozyurda

On behalf of the CMS Collaboration

33rd Deep Inelastic Scattering Workshop
06.05.2026

Monte Carlo Event Generators

- A pp collision includes a primary hard scattering plus additional perturbative and non-perturbative QCD activity, modeled in MC generators via …

  - Matrix Elements/Hard Process

    Fixed order Perturbative QCD/EW
    (LO / NLO / multileg ME)

  - Parton Shower

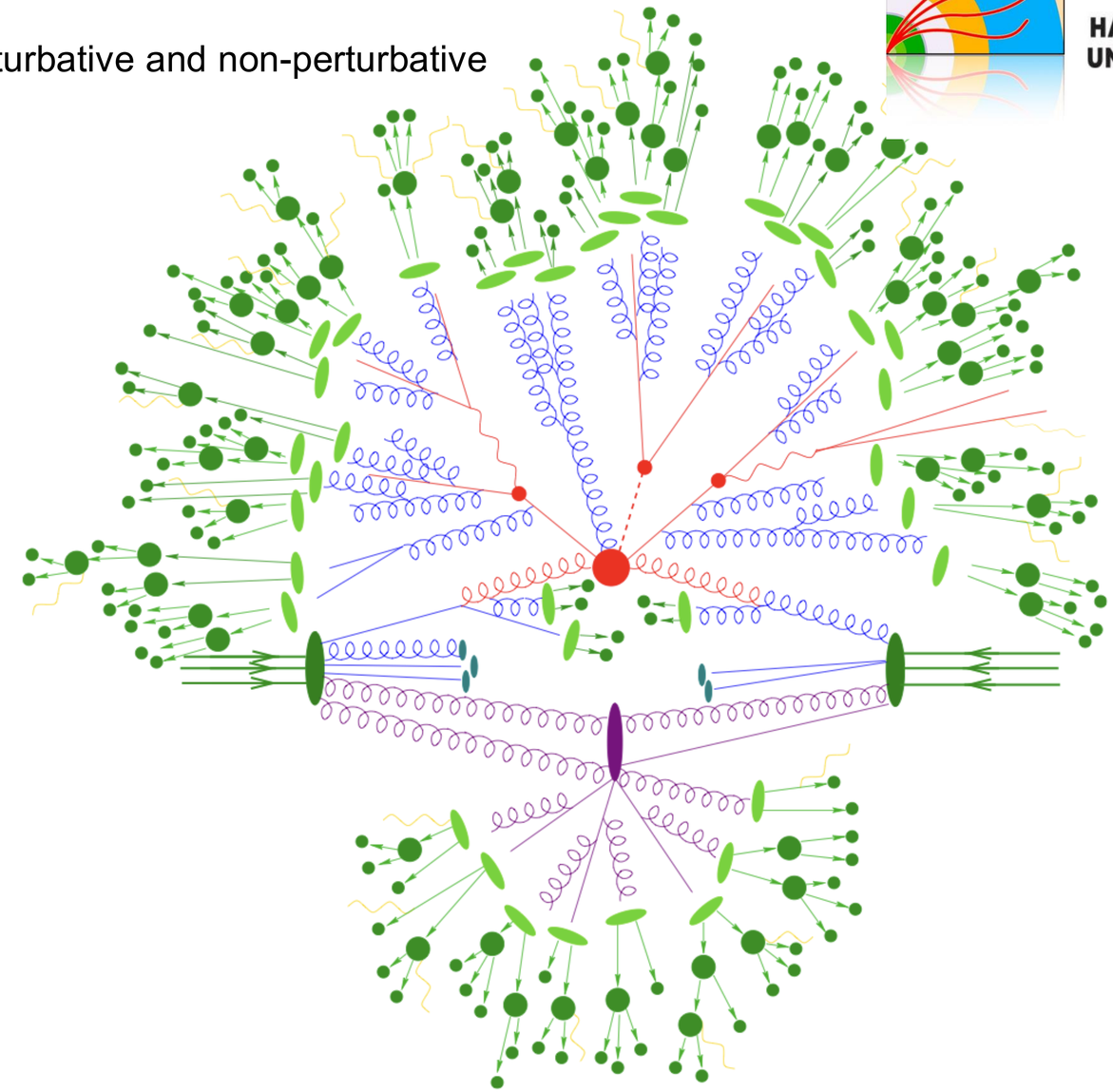
    Soft/collinear QCD radiation
    Leading-log resummation

  - MPI
    Additional semi-hard and soft activity

  - Beam Remnants, Primordial kT

  - Hadronization
    Phenomenological Models (Cluster/String)

  - Hadron Decays
    Effective models constrained by data



F. Krauss

TUNING STUDIES AT CMS

CMS MC Tunes for LHC Run 2 and Run 3

CMS Pythia8 Tunes

- **UE Tunes**
  - **CPX Tunes** (arXiv:1903.12179v2)
  - **CR Dedicated Tunes** (arXiv:2205.02905)
- **Primordial kT Tune** (arXiv:2409.17770)

CMS Herwig7 Tunes

- **UE Tunes**
  - **CHX Tunes** (arXiv:2011.03422)

- *All existing CMS tuning results obtained using Professor (arXiv:0907.2973)*

Fitted Parameters in UE Tunes

$p_{\perp,min} \approx p_{\perp 0}$

The $p_{T,min}$ parameter sets the lower transverse-momentum scale for perturbative parton-parton interactions, thereby regulating the low-p_T divergence of the QCD scattering cross section.

$$\frac{d\hat{\sigma}}{dp_\perp^2} \sim \frac{\alpha_s^2(p_\perp^2)}{p_\perp^4} \longrightarrow \frac{\alpha_s^2(p_{\perp 0}^2 + p_\perp^2)}{(p_{\perp 0}^2 + p_\perp^2)^2}, \qquad p_{T0}(\sqrt{s}) = p_{T0}^{ref}\left(\frac{\sqrt{s}}{s_0}\right)^{\epsilon}.$$

Impact Parameter

The impact parameter describes the overlap between two colliding protons and controls the probability of MPI.

PYTHIA 8

$$\rho(r) \propto \frac{1-\beta}{a_1^3}\exp\left\{-\frac{r^2}{a_1^2}\right\} + \frac{\beta}{a_2^3}\exp\left\{-\frac{r^2}{a_2^2}\right\}$$

a_2: $coreRadius$ and β: $coreFraction$ parameters

HERWIG 7

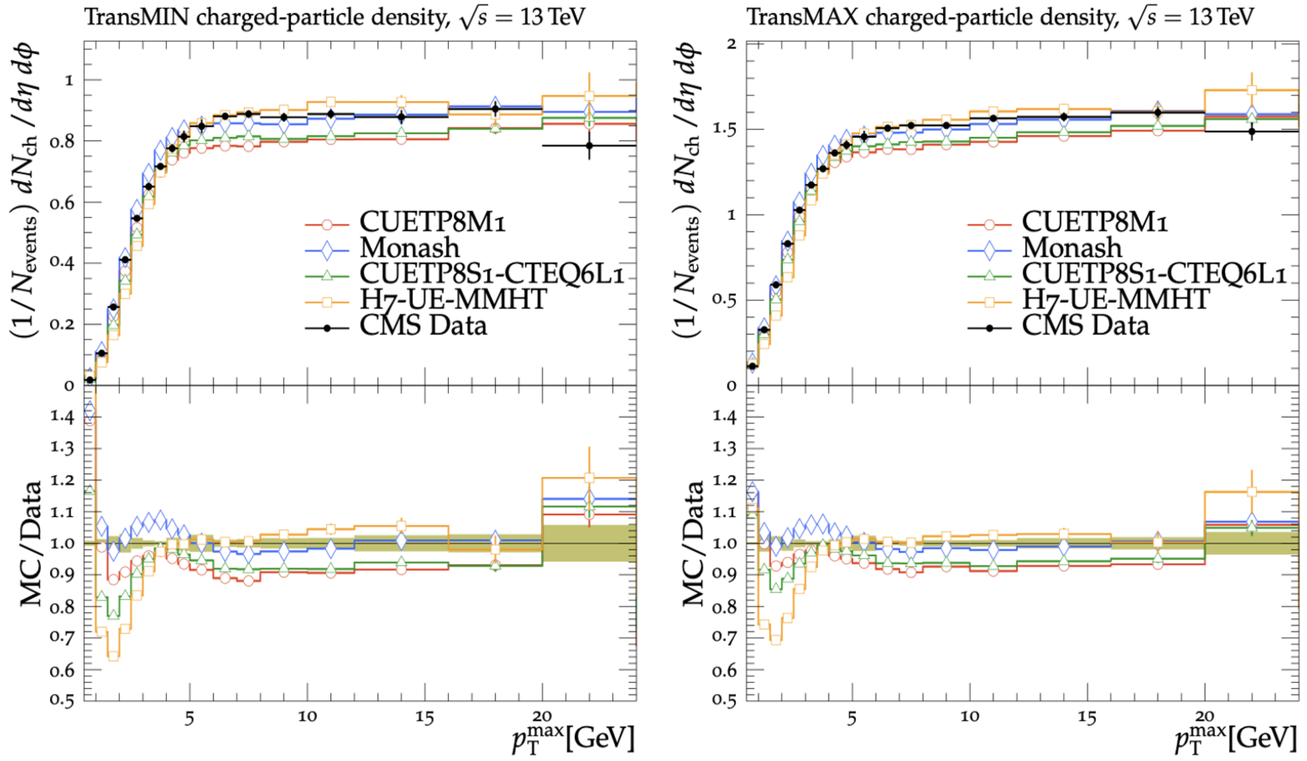
$$G(b) = \int \frac{d^2\mathbf{k}}{(2\pi)^2}\frac{e^{i\mathbf{k}\cdot\mathbf{b}}}{(1+\mathbf{k}^2/\mu^2)^2}.$$

μ^2: $Inverse\ Radius$

Colour Reconnection

PYTHIA 8 uses string-based CR, while HERWIG 7 uses cluster-based CR; the CR probability controls the reconnection strength in the tune.

CPX Tunes: Motivation



arXiv:1903.12179v2

- **Monash 2013** is a general-purpose PYTHIA8 tune based on LEP, Tevatron and early LHC data, constraining shower, hadronization, MPI and UE modelling. (arXiv:1404.5630)
- CUETP8M1:Monash-based tune retuned to CMS UE data (CMS-PAS-TOP-16-021)

MPI and CR parameters retuned; shower settings inherited

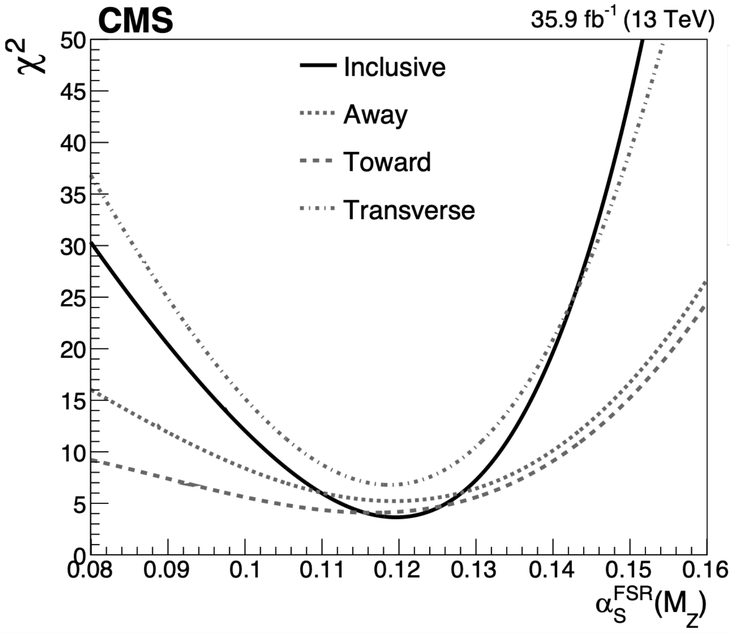
- Inherited Monash shower coupling: $a_s^{ISR} = a_s^{FSR} = 0.1365$

➢ CUETP8M1 failed to describe 13 TeV UE data.

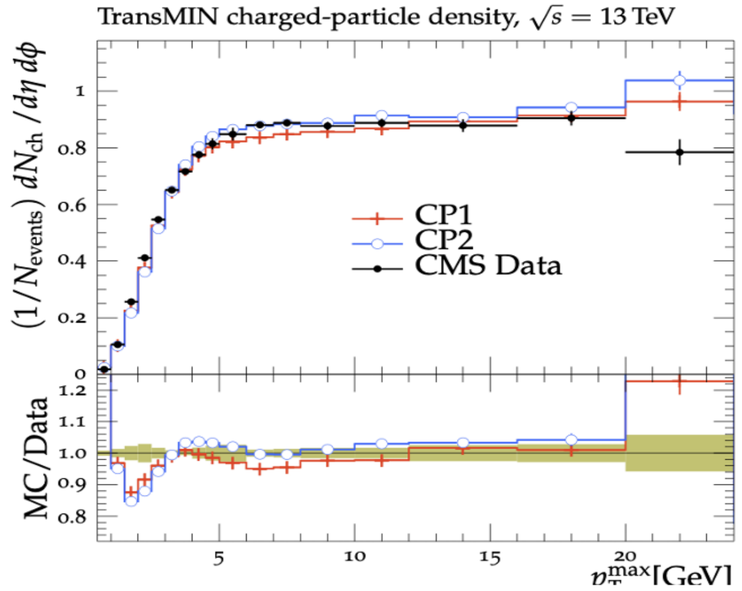
CMS UE in $t\bar{t}$ favors a lower FSR shower coupling:

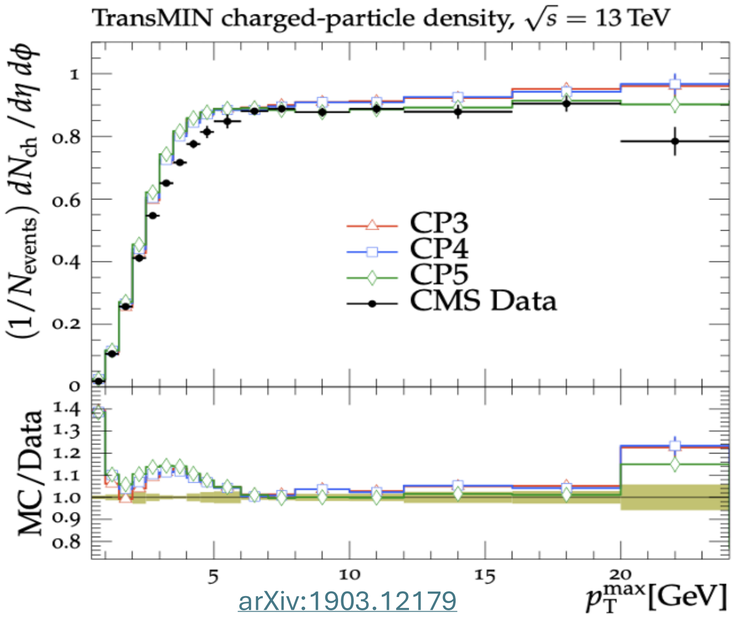
$$\alpha_s^{\mathrm{FSR}}(M_Z) \approx 0.118.$$

a_s^{FSR} *from Underlying Event in ttbar (EPJC 79 (2019) 123)*



CPX TUNES: Results





arXiv:1903.12179

PYTHIA8 parameter	CP1	CP2	CP3	CP4	CP5
PDF Set	NNPDF3.1 LO	NNPDF3.1 LO	NNPDF3.1 NLO	NNPDF3.1 NNLO	NNPDF3.1 NNLO
$\alpha_S(m_Z)$	0.130	0.130	0.118	0.118	0.118
SpaceShower:rapidityOrder	off	off	off	off	on
MultipartonInteractions:EcmRef [GeV]	7000	7000	7000	7000	7000
$\alpha_S^{ISR}(m_Z)$ value/order	0.1365/LO	0.130/LO	0.118/NLO	0.118/NLO	0.118/NLO
$\alpha_S^{FSR}(m_Z)$ value/order	0.1365/LO	0.130/LO	0.118/NLO	0.118/NLO	0.118/NLO
$\alpha_S^{MPI}(m_Z)$ value/order	0.130/LO	0.130/LO	0.118/NLO	0.118/NLO	0.118/NLO
$\alpha_S^{ME}(m_Z)$ value/order	0.130/LO	0.130/LO	0.118/NLO	0.118/NLO	0.118/NLO
MultipartonInteractions:pT0Ref [GeV]	2.4	2.3	1.52	1.48	1.41
MultipartonInteractions:ecmPow	0.15	0.14	0.02	0.02	0.03
MultipartonInteractions:coreRadius	0.54	0.38	0.54	0.60	0.76
MultipartonInteractions:coreFraction	0.68	0.33	0.39	0.30	0.63
ColorReconnection:range	2.63	2.32	4.73	5.61	5.18
χ^2/dof	0.89	0.54	0.76	0.80	1.04

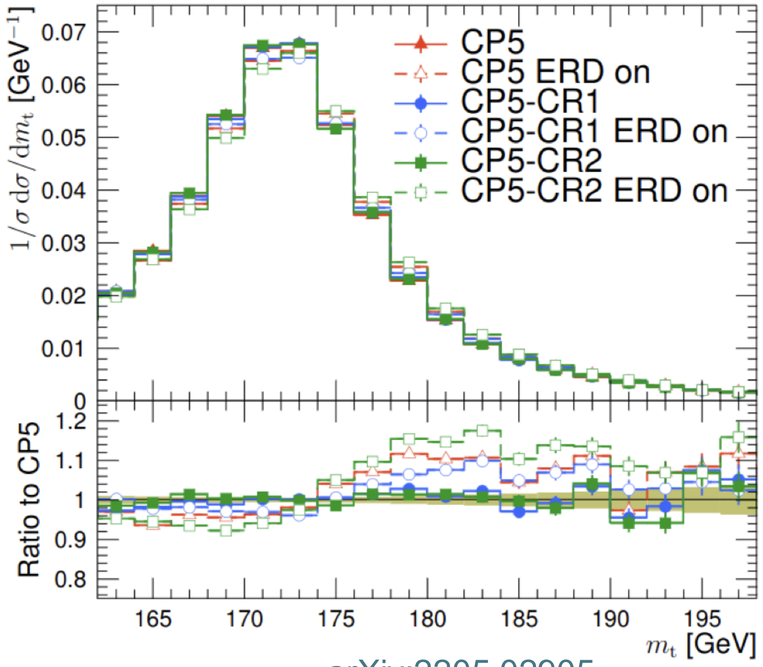
- CP tune extraction performed with PYTHIA 8.226 (arXiv:1410.3012)
- CP1–CP5 represent distinct PYTHIA 8 configurations with different PDF orders, α_s choices, and ISR settings.
- CP5: best UE agreement; validated across a wide range of final states (UE, MB, top, V+jets)
- CP5 was therefore recommended by CMS as the main tune for Run 2 and Run 3, and it is still in use.

Colour Reconnection Dedicated Tunes

Table 2: The MPI and CR parameter ranges used in the tuning procedure.

PYTHIA 8 parameter	Min–Max
MPI parameters	
`MultipartonInteractions:pT0Ref`	1.0—3.0
`MultipartonInteractions:ecmPow`	0.0—0.3
`MultipartonInteractions:coreRadius`	0.2—0.8
`MultipartonInteractions:coreFraction`	0.2—0.8
QCD-inspired model	
`ColourReconnection:m0`	0.1—4.0
`ColourReconnection:junctionCorrection`	0.01—10
`ColourReconnection:timeDilationPar`	0—60
Gluon-move model	
`ColourReconnection:m2lambda`	0.2—8.0
`ColourReconnection:fracGluon`	0.8—1.0

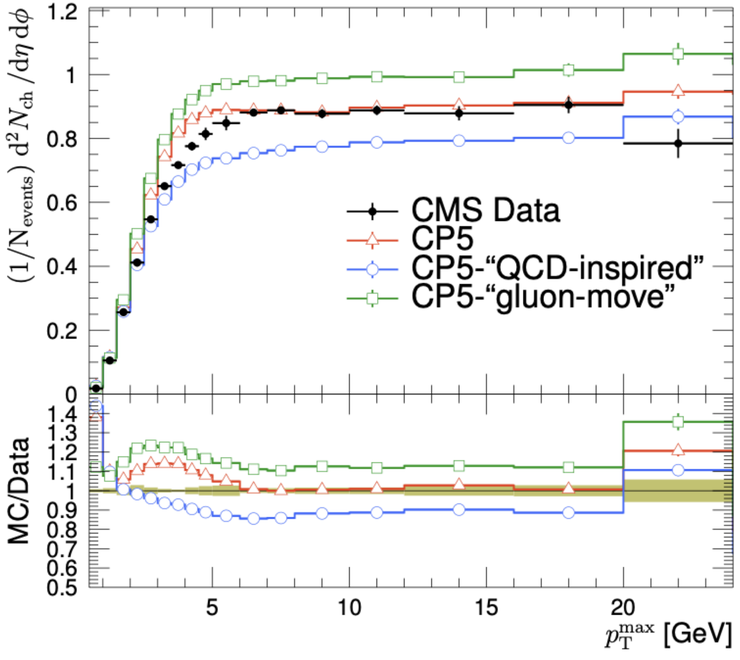
Invariant mass of top quark candidates, $\sqrt{s} = 13$ TeV



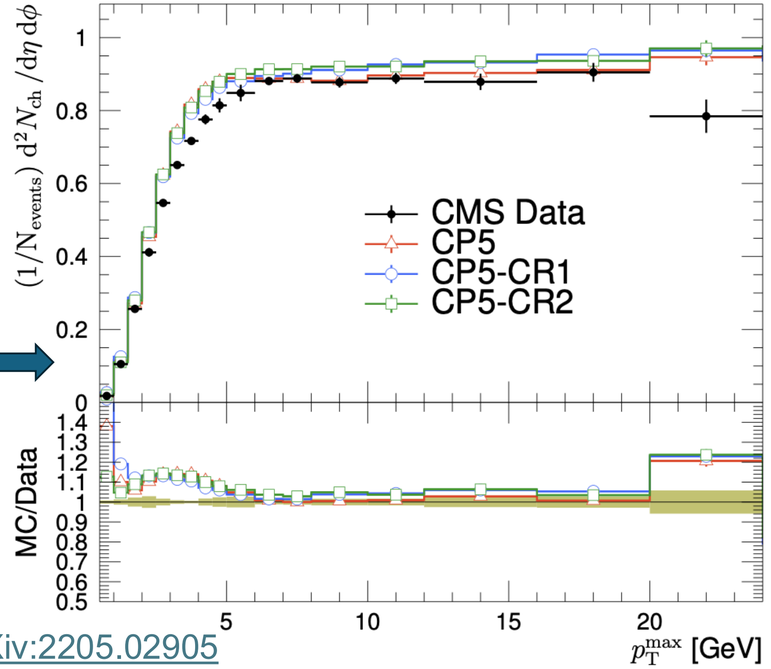
arXiv:2205.02905

- CP5 uses the default **MPI-based** colour reconnection model,
- QCD-inspired and gluon-move alternatives were tested as dedicated CR variations.

TransMIN charged-particle density $\sqrt{s} = 13$ TeV



TransMIN charged-particle density $\sqrt{s} = 13$ TeV



arXiv:2205.02905

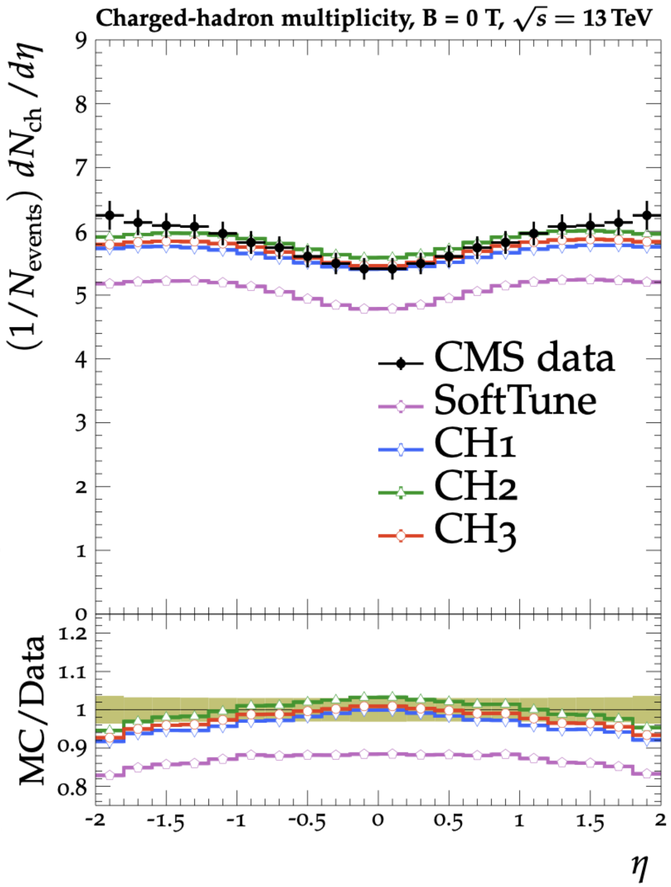
- Alternative CR models, once tuned, describe the data at a similar level as the MPI-based model

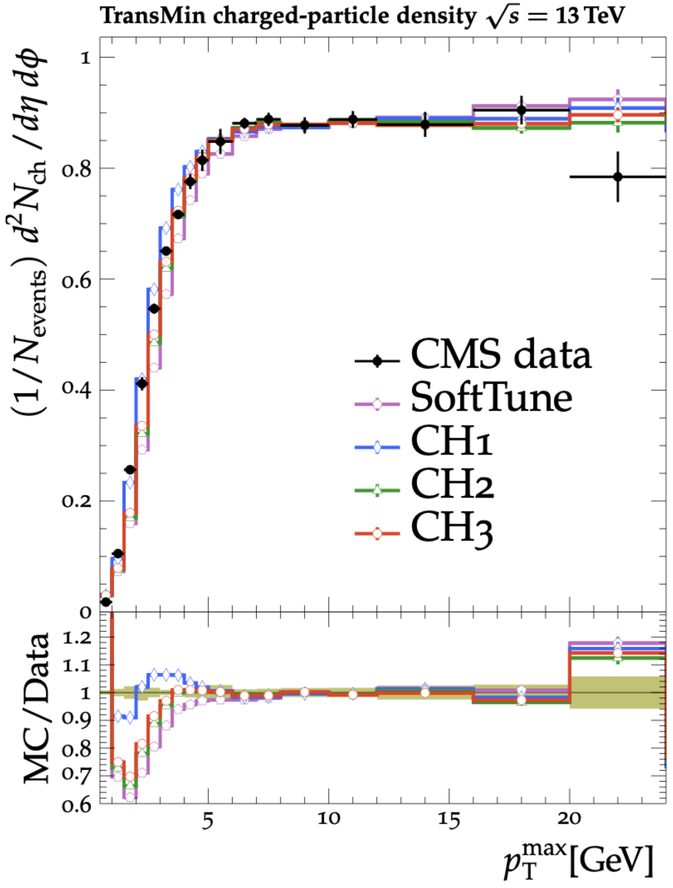
CR tune variations provide a more realistic estimate of the modelling uncertainty in m_{top} than a simple CR on/off approach.

CHX TUNES

Table 2: Value of the parameters for the SoftTune [3, 12], CH1, CH2, and CH3 tunes.

		SoftTune	CH1	CH2	CH3
$\alpha_S(m_Z)$		0.1262	0.118	0.118	0.118
PS	PDF set	MMHT 2014 LO	NNPDF 3.1 NNLO	NNPDF 3.1 NNLO	NNPDF 3.1 NNLO
	$\alpha_S^{PDF}(m_Z)$	0.135	0.118	0.118	0.118
MPI & remnants	PDF set	MMHT 2014 LO	NNPDF 3.1 NNLO	NNPDF 3.1 LO	NNPDF 3.1 LO
	$\alpha_S^{PDF}(m_Z)$	0.135	0.118	0.118	0.130
$p_{\perp,0}^{min}$ (GeV)		3.502	2.322	3.138	3.040
b		0.416	0.157	0.120	0.136
μ^2 (GeV^{-2})		1.402	1.532	1.174	1.284
p_{reco}		0.5	0.400	0.479	0.471
χ^2/N_{dof}		12.8	6.75	1.54	1.71





arXiv:2011.03422

- SoftTune: previous HERWIG 7.1 MPI tune, based on MMHT2014 LO PDFs and ATLAS MB data.

- CH1–CH3 are CMS HERWIG 7 UE tunes obtained with HERWIG v7.1.4. (arXiv:1705.06919)

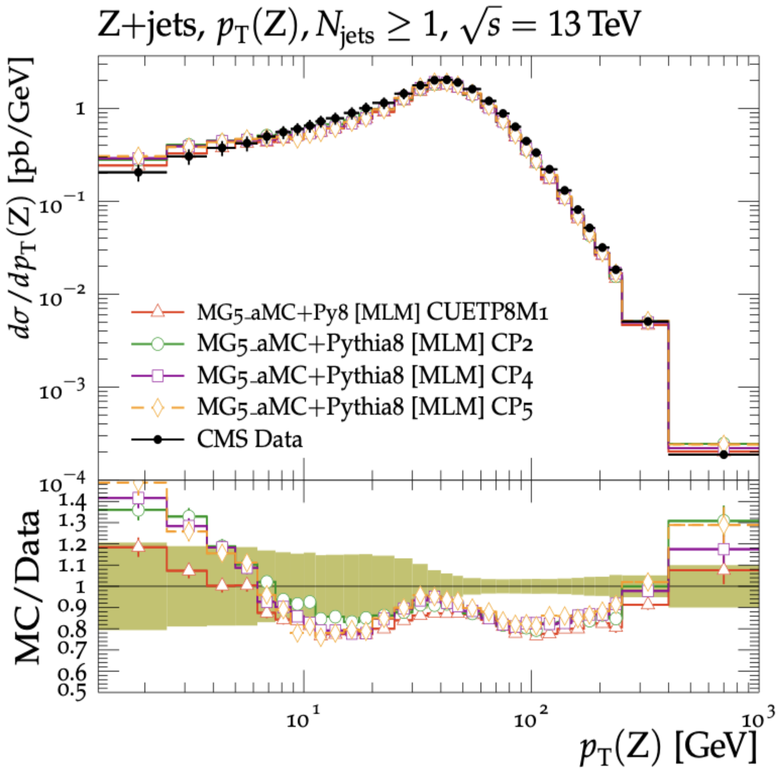
- They use CMS MB and UE measurements with NNPDF3.1-based PS and MPI/beam-remnant setups.

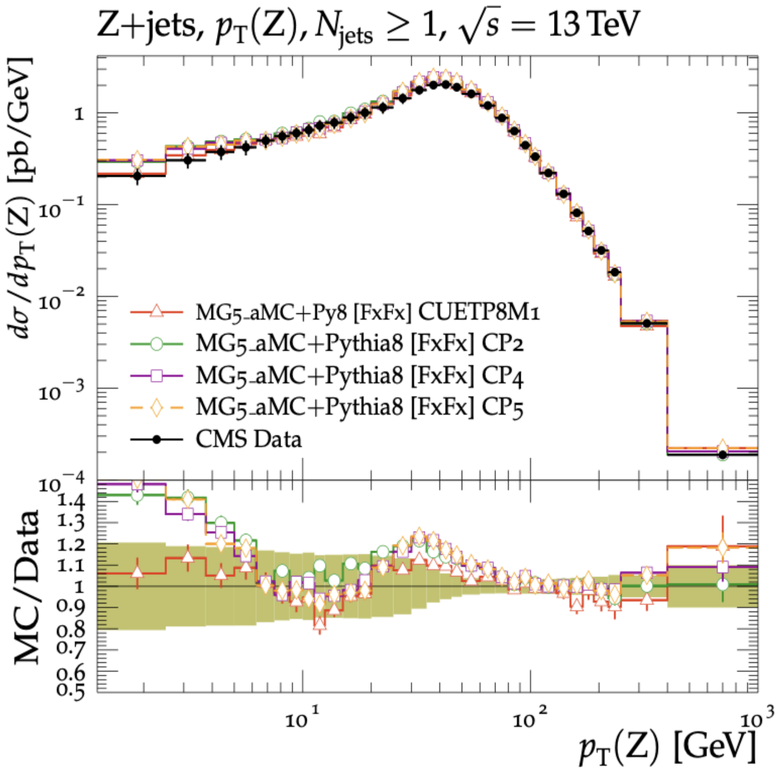
- CH3 is the preferred tune, improving the UE description relative to SoftTune and CH1.

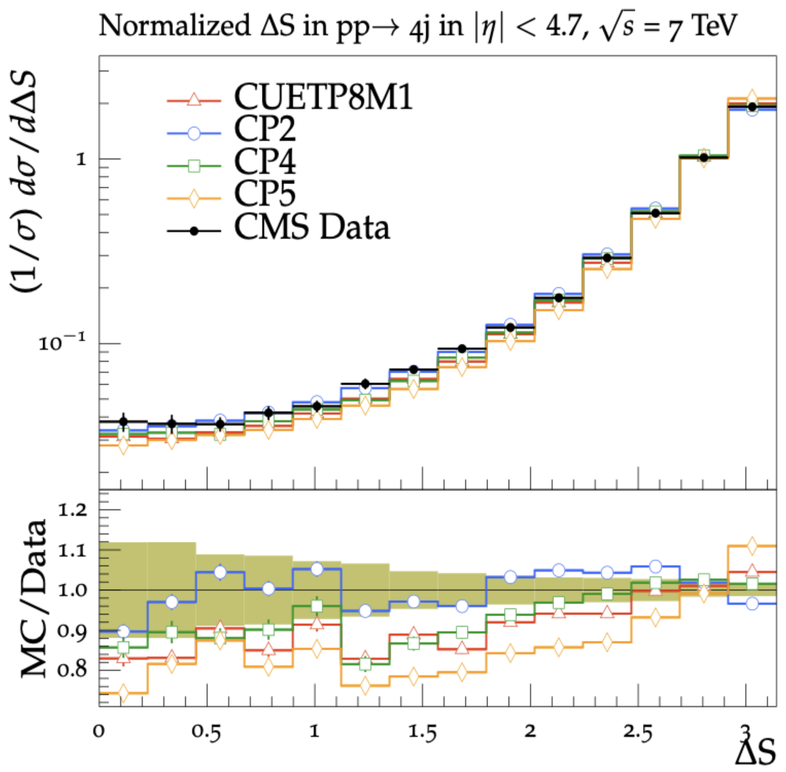
Tensions of CH3 and CP5 Tunes

- Not all tensions can be removed by tuning
- Tuning procedure can only optimize an existing model within its theoretical and parametric limitations.
- Despite their overall strong performance, CP5 and CH3 still exhibit tensions in specific observables (*e.g., DY+jets, double parton scattering, and jet substructure*).

arXiv:1903.12179

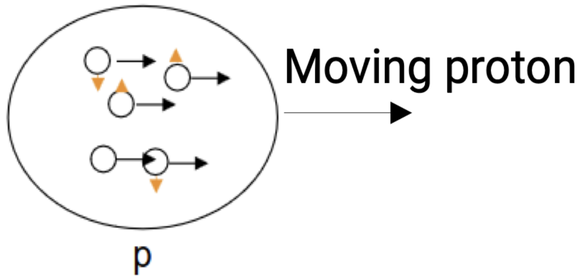






Primordial kT

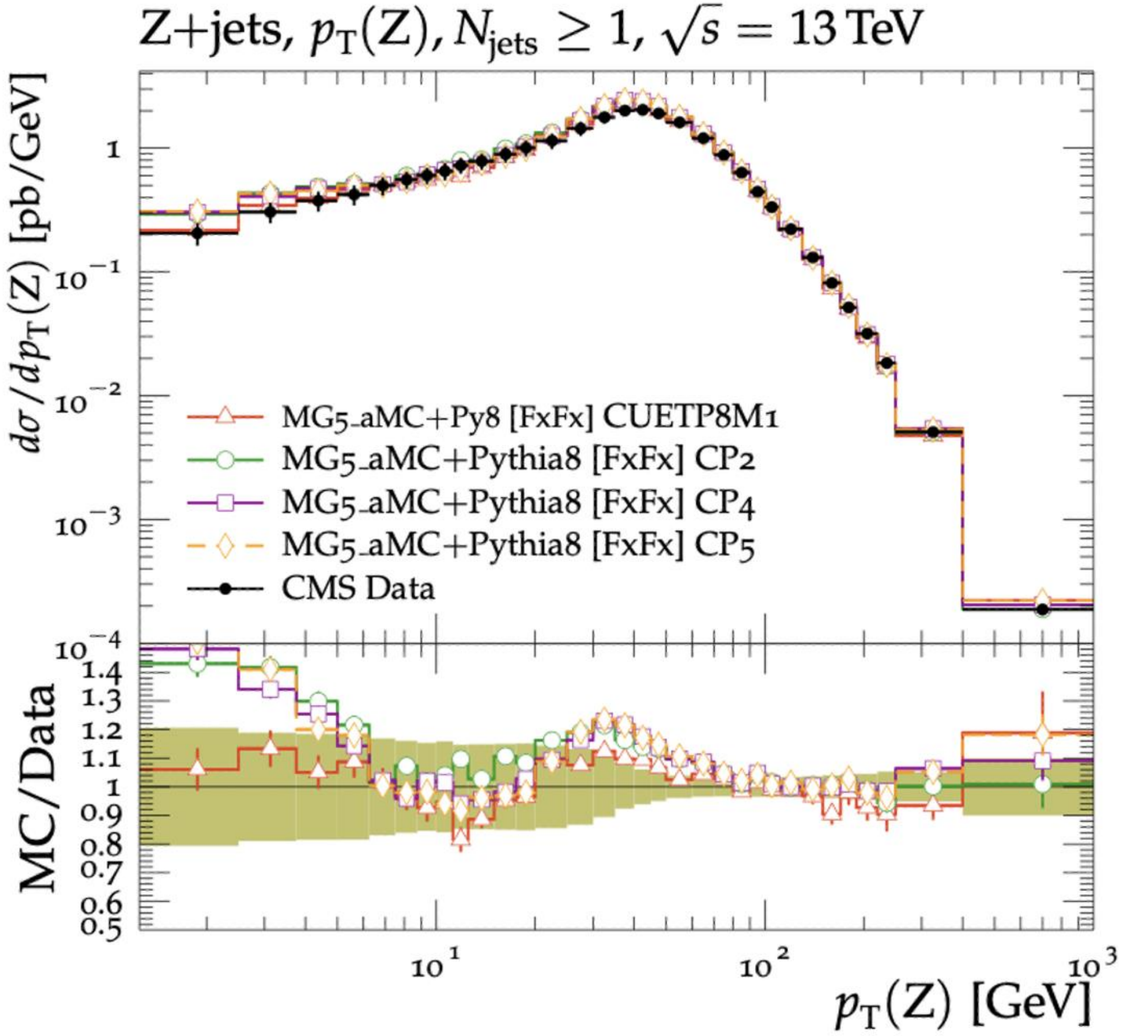
- Primordial kT represents the non-zero intrinsic transverse momentum of partons inside the proton.
- At low pT, interplay with radiation e.g. ISR.
- It accounts for non-perturbative effects
- In Monte Carlo generators, it is implemented through an effective Gaussian smearing model



$$\sigma(Q) = \frac{\sigma_{\text{soft}} Q_{1/2} + \sigma_{\text{hard}}}{(Q_{1/2} + Q)} \frac{m}{(m + m_{1/2} y_{\text{damp}})},$$

The width of this distribution, $\sigma(Q)$,controls the typical transverse momentum scale

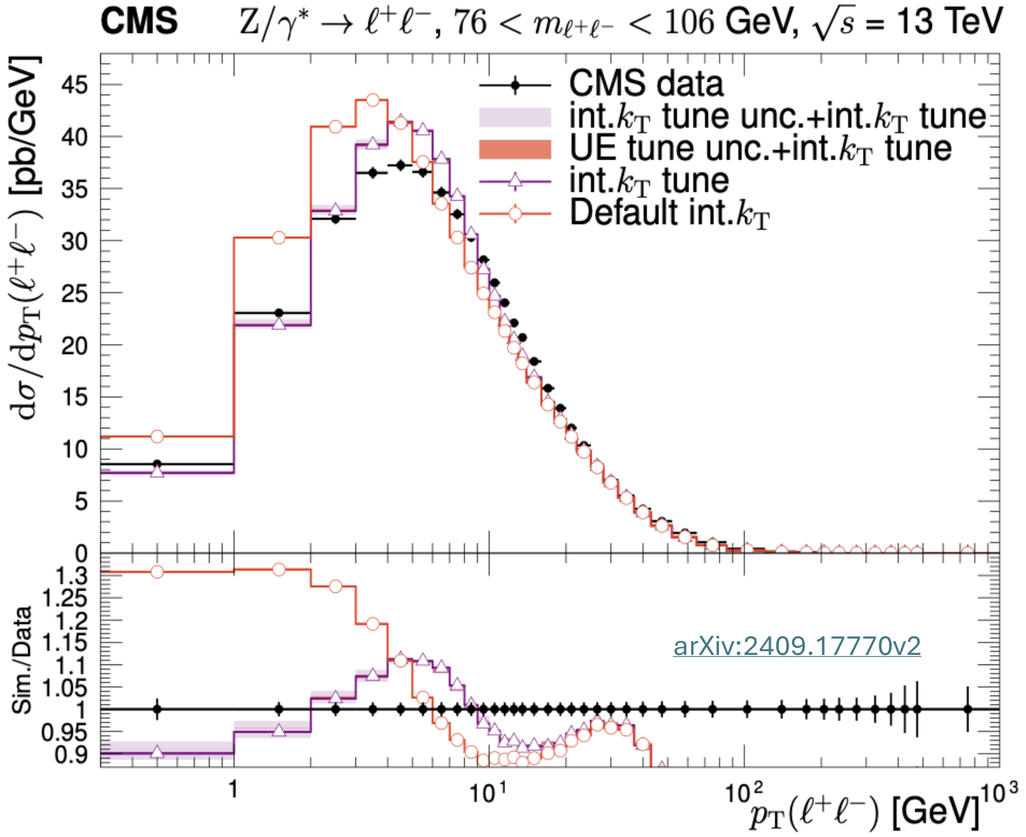
arXiv:1903.12179



Primordial kT Tune

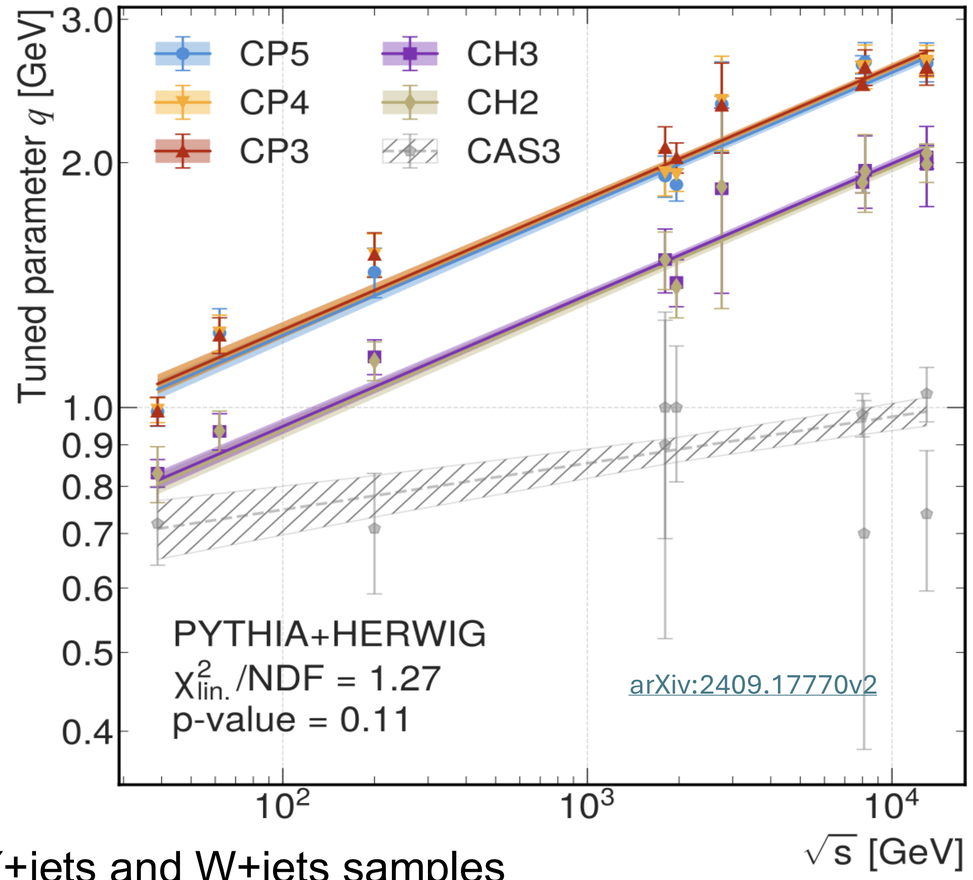
• Intrinsic kT is extracted from low-pT Drell–Yan spectra.

• The tuned samples improve the data–MC agreement in the low-pT region.



• Primordial kT is √s dependent,

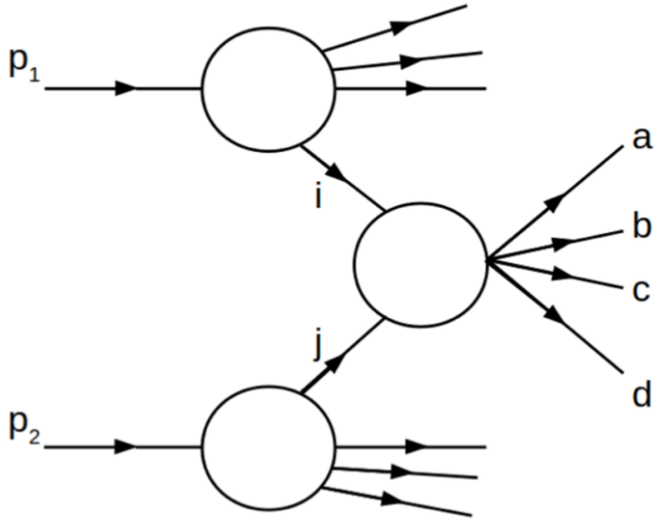
• The $\sqrt{s}$ scaling is similar for PYTHIA and HERWIG, while the overall normalization differs.



➢ **CP5-kT** tune used in Run 3 DY+jets and W+jets samples

ONGOING TUNING STUDIES AT CMS

Double Parton Scattering (DPS)



SINGLE PARTON SCATTERING (SPS)

- One hard scattering in a single proton proton collision. Final states particles are correlated.
- **correlated topologies**, back-to-back jets

DPS sensitive observables:

- The difference in azimuthal angle between the light jet pair

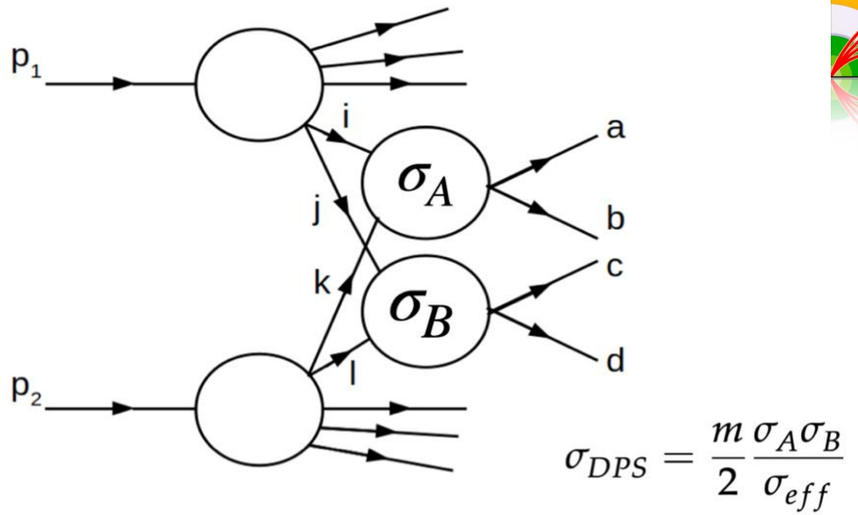
$$\Delta\phi = |\phi(j_1) - \phi(j_2)|$$

- The balance in p_T of the two light jets:

$$\Delta p_T = \frac{|p_T(j_1) + p_T(j_2)|}{(||p_T(j_1)| + |p_T(j_2)||)}$$

- The azimuthal angle between the two dijet pairs:

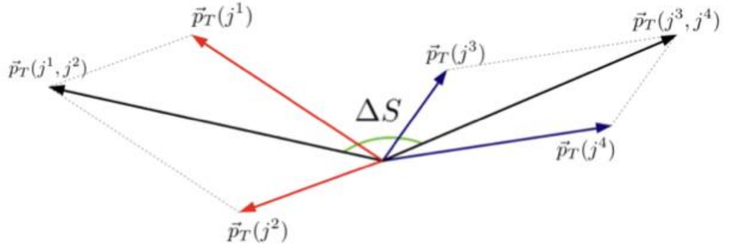
$$\Delta S = \frac{(p_T(j_3, j_4).p_T(j_1, j_2))}{(|p_T(j_3, j_4)| + |p_T(j_1, j_2)|)}$$



$$\sigma_{DPS} = \frac{m}{2}\frac{\sigma_A \sigma_B}{\sigma_{eff}}$$

DOUBLE PARTON SCATTERING (DPS)

- Two **separate** hard interactions in a single proton-proton collision. Two pairs of partons from the incoming hadrons interact **independently** with each other.
- **Uncorrelated topologies**, back-to-back jets only for each of the independently produced jet pairs



Why DPS Processes are Important?

DPS cross section is suppressed w.r.t SPS:

$$\frac{\sigma_{DPS}}{\sigma_{SPS}} \sim \frac{\Lambda^2}{Q^2}$$

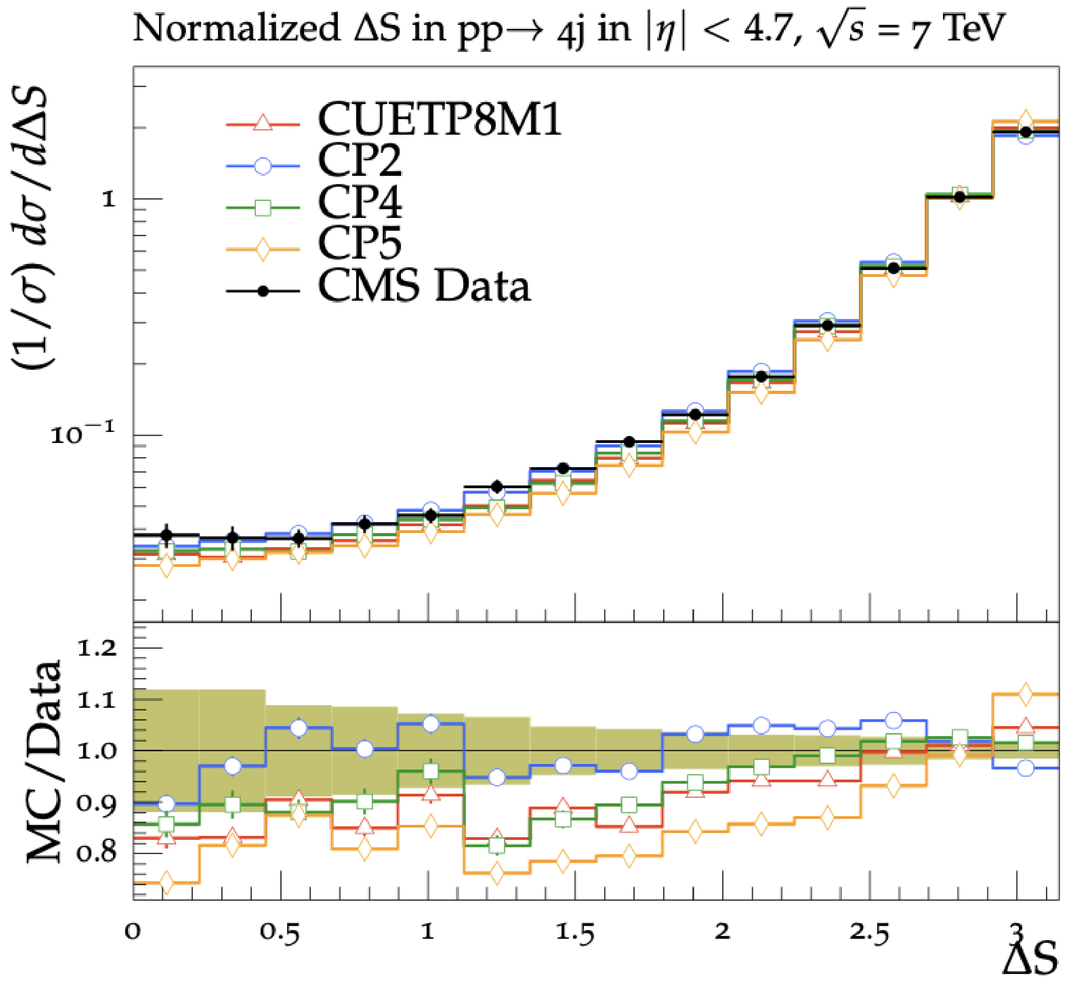
(Hadronic scale ~1GeV)

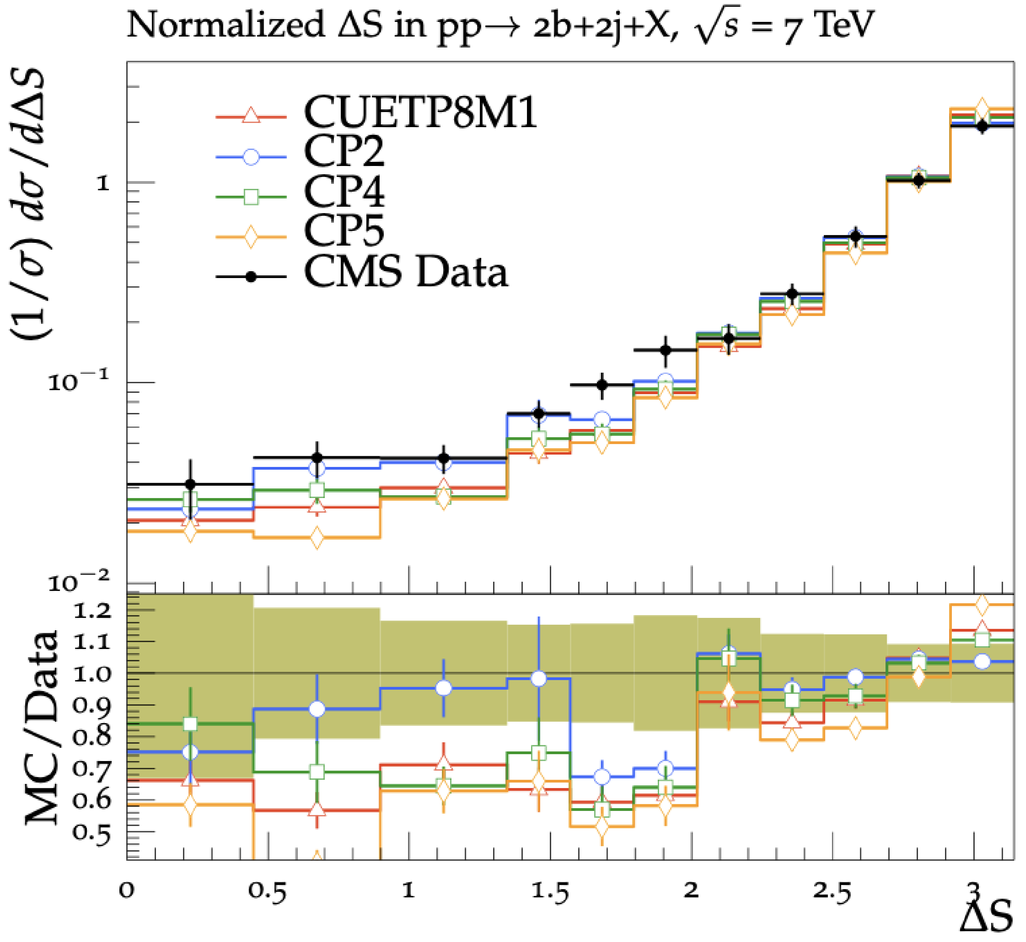
(Hard interaction scale)

- DPS becomes more important as the collider energy grows. Larger density of partons at small-x values.
- DPS provides a unique opportunity to probe the internal structure of hadrons, particularly protons.
- DPS can contribute to the background in various analyses; Higgs, BSM, top, high-multiplicity final states.
- DPS can become competitive with SPS when SPS is hindered by small couplings; e.g. same-sign WW production.

Motivation of DPS Tune: Plots

arXiv:1903.12179

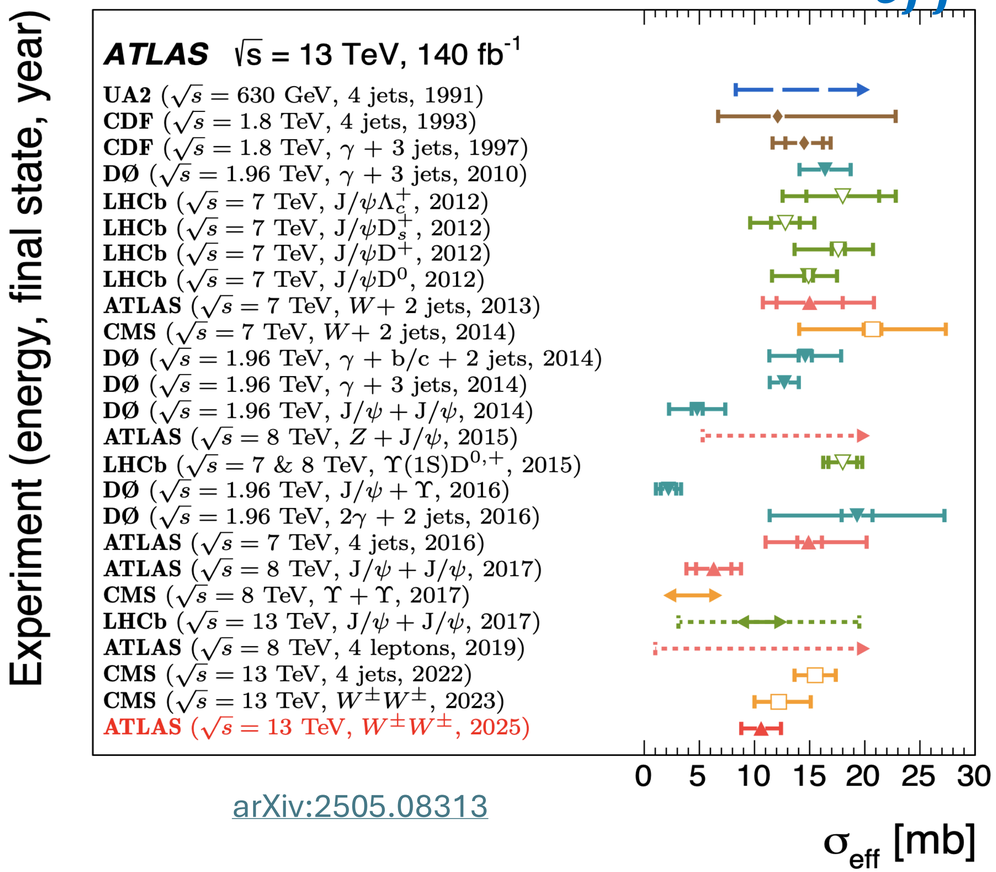




- CP5 and CH3 shows tensions in DPS-sensitive observables.

*CUETP8M1:(CMS-PAS-TOP-16-021)

Motivation of DPS Tune: σ_{eff}



arXiv:2505.08313

	$\sqrt{s} = 7\,\text{TeV}$ σ_{eff} [mb]	$\sqrt{s} = 13\,\text{TeV}$ σ_{eff} [mb]
CP1	$26.3^{+1.0}_{-1.7}$	$27.8^{+1.1}_{-1.4}$
CP2	$24.7^{+1.0}_{-1.6}$	$26.0^{+1.0}_{-1.3}$
CP3	$24.1^{+1.0}_{-1.5}$	$25.2^{+1.0}_{-1.3}$
CP4	$23.9^{+1.0}_{-1.5}$	$25.3^{+1.1}_{-1.4}$
CP5	$24.0^{+1.0}_{-1.6}$	$25.3^{+1.0}_{-1.3}$

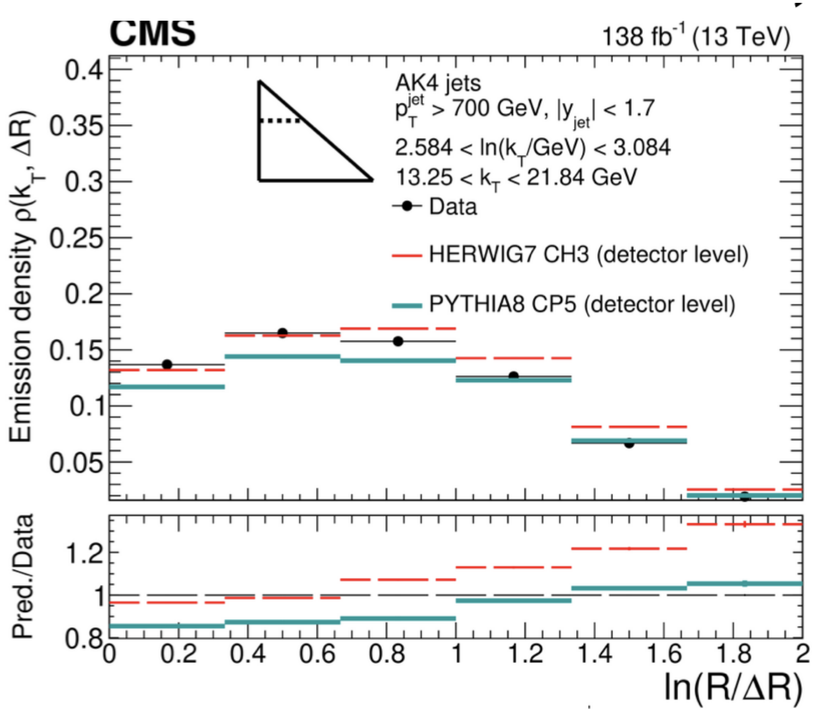
$$\sigma_{eff} = \frac{1}{\int d^2b\, F^2(b)} = \frac{\sigma_{ND}}{f_{EF}} \qquad \sigma_{DPS} = \frac{m}{2}\frac{\sigma_A \sigma_B}{\sigma_{eff}}$$

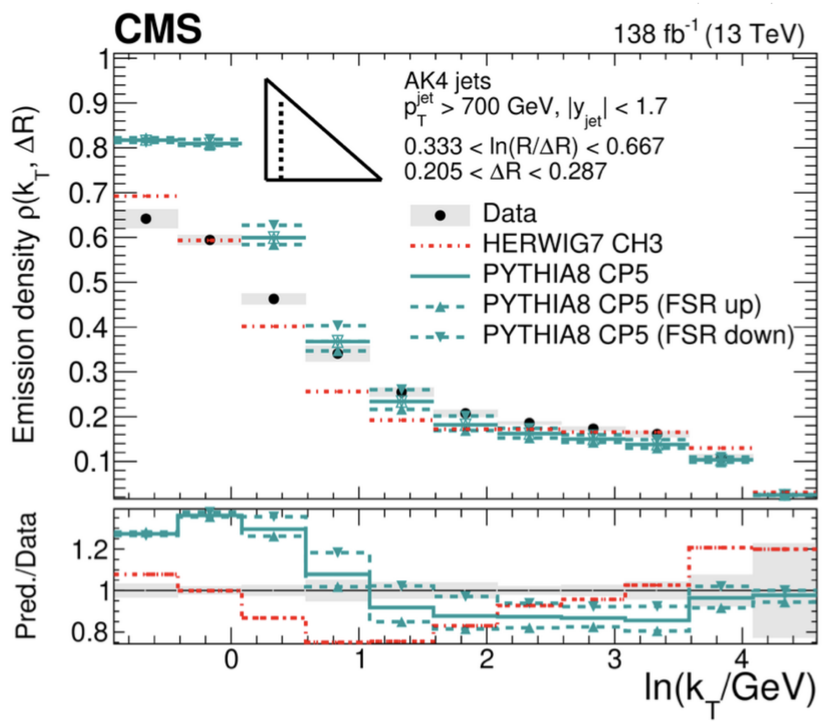
- Differences between σ_{eff} measurements and MC predictions motivate improved DPS modelling.

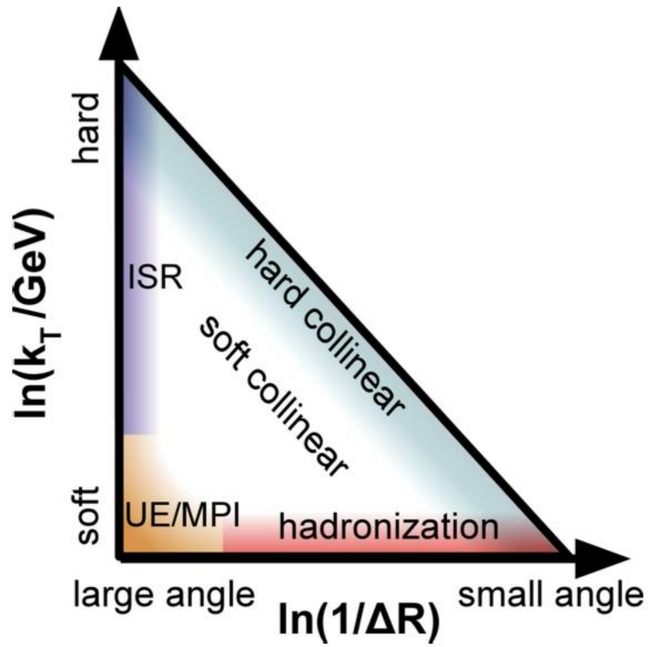
CMS 13 TeV 4-jet result is $\sigma_{eff} = 15.5 \pm 1.8$ mb and
ATLAS 7 TeV 4-jet measurement is reported as $\sigma_{eff} = 14.9^{+1.2}_{-1.0}\,(stat)^{+5.1}_{-3.8}\,(syst.)mb$

- Starting from CMS UE tunes, we investigate different CR and beam-profile models to improve the description of DPS-sensitive multijet measurements.

Towards jet substructure tune







SMP-22-007: arXiv:2312.16343

- The **Lund Jet Plane** separates emissions by angular scale and relative hardness
- SMP-22-007 shows sizeable mismodelling for CP5 and CH3 in both perturbative and transition regions
- These deviations indicate that **FSR and hadronization settings need dedicated constraints**
- First tuning step: vary $\alpha_s^{\mathrm{FSR}}(m_Z)$,then extend to FSR cutoff, Lund fragmentation parameters, CR models and complementary jet-substructure observables

Summary

- Event-generator tuning is essential for reliable MC predictions
- CP5 and CH3 remain the main CMS reference tunes for Run 2 and Run 3 studies
- Some modelling limitations have been identified and addressed through dedicated tuning efforts
- Intrinsic k_T tuning improves the modelling of initial-state transverse momentum effects
- CP5-k_T is already used in Run 3 DY+jets and W+jets samples
- Ongoing DPS and LJP studies target measurements where CP5 and CH3 show tensions, aiming for improved dedicated tunes

THANKS

BACK-UP

Underlying Event

- In hadron-hadron collisions, the **underlying event** denotes the activity not directly associated with the primary hard scattering

- And is mainly driven by multiple parton interactions, beam remnants, and color reconnection, with additional contributions from parton showering and hadronization.

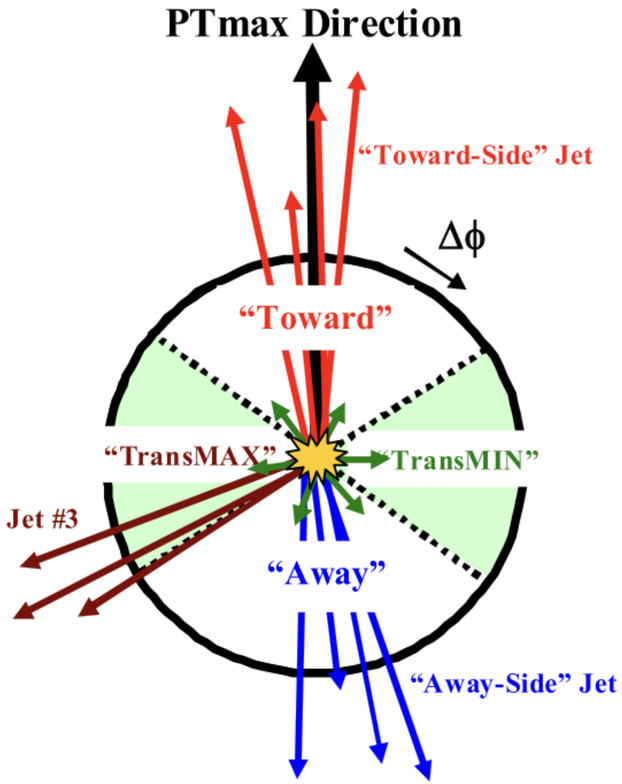
- The transverse region provides the most useful operational definition of the underlying event

- As it is less directly populated by the leading hard scatter and its recoil

Transverse: $60 < |\Delta\varphi| < 120$

- TransMax: maximum activity side <== MPI/BR & ISR/FSR
- TransMin: minimum activity side <== MPI/BR
- TransDIF: TransMax - TransMin

The phenomenological parameters of MC event generators must be constrained through **tuning** to experimental data.

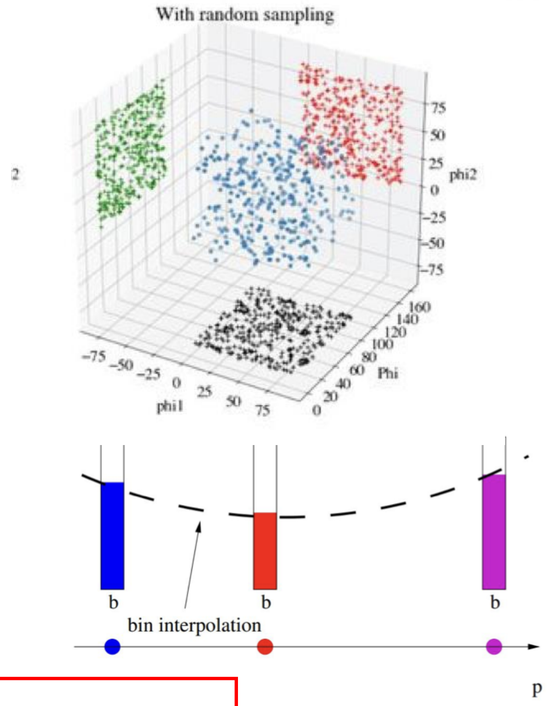


Tuning Framework: Professor

- To avoid the high cost of direct MC optimisation, **Professor** replaces the generator response with a fast surrogate model based on **polynomial interpolation**.

Tuning Flow

- For the targeted tuning, minimum and maximum ranges are defined for each free parameter of the selected phenomenological model(s), taking into account physical expectations and model limitations.
- The parameter space is then sampled randomly within the predefined ranges.
- From these samples, a fast surrogate/interpolation model is built for each observable bin.
- The best-fit parameters are then determined by minimizing χ^2on the surrogate model instead of the full MC prediction.



$$\chi^2(p) = \sum_O w_O \sum_{i \in O} \frac{(f^i(p) - R_i)^2}{\Delta_i^2}$$

p: tuning parameters
$\mathcal{O}$**: binned observable**
w_O **:observable weight**
$f^i(p)$ **:parametrized prediction for bin i**
R_i **:measured content of bin i**
Δ_i **:uncertainty in bin** i

Limitations

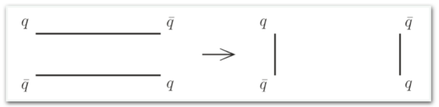
- Limitations over number of tuning parmeters
- The heuristic χ^2 treatment is limited by the absence of bin-to-bin correlations.
- Observable weighting must be assigned manually.

Alternative CR Models in Pythia 8

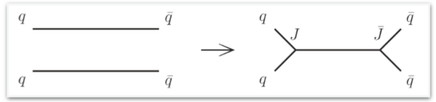
➢ Beyond the MPI-based CR model used in CP5, PYTHIA8 includes two alternative colour reconnection schemes.

QCD-inspired colour reconnection model (CP5-CR1)

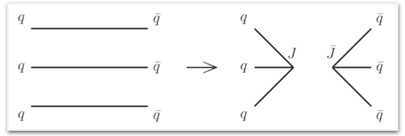
- Builds on the usual string-length minimization picture, but additionally requires that any reconnection be allowed by QCD colour rules.
- In this way, the model selects colour rearrangements that are both dynamically favoured and colour compatible, and it can also form **junction structures**, which improves the description of baryon production.



ordinary string reconnection



double junction reconnection

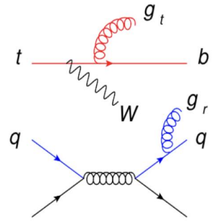


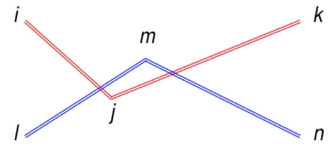
triple junction reconnection

more types of junction reconnections available..

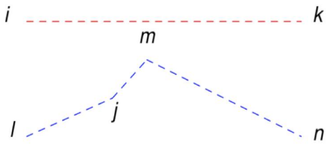
Gluon-move colour reconnection model (CP5-CR2):

- Based on an iterative procedure in which final-state gluons are reassigned between colour-connected string systems whenever this leads to a reduction in the total string length.
- The model therefore modifies the colour topology by relocating gluons to more favourable string configurations.





original configuration



resulting configuration after moving the gluon (reduced λ)

$$\Delta\lambda(j, lm) = \lambda_{j;lm} - \lambda_{j;ik} = \lambda_{lj} + \lambda_{jm} + \lambda_{ik} - (\lambda_{ij} + \lambda_{jk} + \lambda_{lm}) .$$

$$\min_{j,lm} \Delta\lambda(j, lm) \leq \Delta\lambda_{\text{cut}}$$

➢ Compared with the default MPI-based CR model, both approaches provide a more structured treatment of the colour topology:

➢ CP5-CR1 emphasizes QCD-motivated allowed reconnections, while CP5-CR2 emphasizes string rearrangement through gluon relocation.

- Christiansen, J.R. & Skands, P.Z. J. High Energ. Phys. (2015) 2015: 3.
- 10.1007/JHEP11(2014)043