Event shape variables in pp collisions in CMS

Soumyadip Barman

Visva-Bharati, India

On behalf of the CMS Collaboration

EPS-HEP Conference

Marseille, France

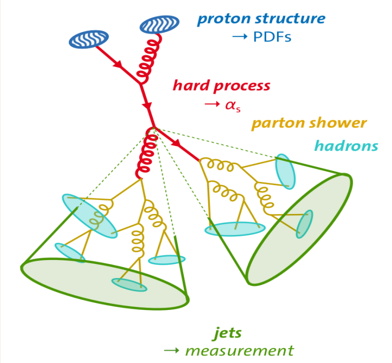
July 9, 2025

Outline

- **QCD at the LHC**
- **Why study ESVs?**
- **ESV Measurements**
- **Recent Studies on ESV in CMS**
  - ***Measurement of event shape variables using charged particles inside jets in proton-proton collisions at $\sqrt{s} = 13$ TeV***
  - ***Measurement of event shapes in minimum bias events from pp collisions at 13 TeV***
- **Conclusions**

QCD at the LHC

- **Hard and Soft QCD** → dominant production methods at the LHC
- Crucial for understanding → **proton structure & strong interaction**
- All LHC observations → depend on the **modelling of the QCD**
- Provides essential experimental input for improving theoretical models → **parton showers, hadronization etc.**



- Several Jet measurements in CMS, studied for understanding QCD –
  - **Event Shape Variables (ESV)**
  - **Differential Jet Cross Sections**
  - **Primary Lund Jet Plane Density**
  - **Energy Correlators**
  - **Jet Charge**
  - . . .

Why study ESVs ?

Event shape variables (ESVs) → defined as **ratios of momenta of the final state objects** (particles, jets ...)

Sensitive to various aspects of QCD –

- **Hard scattering part** – analytically calculable with good accuracy
  - ***Complement of Transverse Thrust*** $(\tau_\perp)$
  - ***Third-Jet Resolution Parameter*** $(Y_{2,3})$
  - ***Sphericity*** (S), ***Transverse Sphericity*** $(S_\perp)$
- **Hadronization part** – described by phenomenological models
  - ***Jet broadening*** (B_T)
  - ***Total jet mass***(ρ_{Tot}), ***Total transverse jet mass*** (ρ^T_{Tot})
  - ***Particle multiplicity and*** p_T ***in event, jets***

- Sensitive → **geometrical properties of energy flow**
- **collinear and infrared divergences** → robust and safe

- Discriminate → **SM** & **New Physics**
- Tests → **parton shower** & **hadronization models**
- **Tune** → MC event generators

ESV Measurements

Before LHC

- Determination of the α_s : In e^+e^- **LEP experiments** **[J. Phys. G 30 (2004)]**
- Studied in **Deep Inelastic** (DIS) ep collisions at **HERA** **[Eur.Phys.J.C 46, 343-456 (2006)]**
- Measurement in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV **[Phys. Rev. D 83, 112007]**

ESVs at LHC

- **CMS** made the first study of the ESVs at the LHC with **jets as inputs** – **[Phys. Lett. B 699 (2011)]**
- Several studies of different ESVs have been done at **CMS** and **ATLAS** with jets as inputs at $\sqrt{(s)} =$ **7 TeV and 13 TeV** – **[JHEP12(2018)117]**, **[JHEP01(2021)188]**
- **ATLAS** and **ALICE**, have also examined ESVs with **charged particles** – **[PhysRevD.88.032004]**, **[Eur.Phys.J.C 72 (2012)]**

Recent study of ESVs in CMS → This Talk

- *Measurement of event shape variables using charged particles inside jets in proton-proton collisions at* $\sqrt{s} = 13$ *TeV* **[CMS-PAS SMP-22-004]**
- *Measurement of event shapes in minimum bias events from pp collisions at 13 TeV* **[CMS-PAS SMP-23-008]**

Measurement of event shape variables using charged particles inside jets in proton-proton collisions at $\sqrt{s} = 13$ TeV

CMS-PAS SMP-22-004

Introduction & Motivation

- Five ESVs are calculated – ***Complement of Transverse Thrust ($\tau_{\perp}$)***, ***Third-jet Resolution Parameter ($Y_{2,3}$)***, ***Total Jet Broadening (B_T)***, ***Total Jet Mass (ρ_{Tot})***, ***Total Transverse Jet Mass(ρ^T_{Tot})***
- Compared with MC predictions → **PYTHIA8**, **HERWIG7**, **MADGRAPH5_AMC@NLO**
- Finding the scope for → **MC tuning** & **theoretical understanding**

Charged particles as inputs

- Improved **event reconstruction** and **particle identification(PF)** → possible to have **particle information in jets with good amount of details**
- Reduces **globalness** issues, offers a finer view of **shower development** and **energy flow**
- ESVs are safe from → **PU** and **UE** effect

Data & Event Selection

- **Data** : **Full Run-2** – $\mathcal{L} = 138 fb^{-1}$
- **Events Selection** –
  - $N_{Jets} \geq 2$
  - **Jets** → $p_T > 30$ GeV & $|\eta| < 2.1$
  - **Charged particles** → $|\eta| < 2.5$
- Data sample divided into **8** $H_{T,2}$ ranges based on **trigger efficiency**
  $H_{T,2} = \frac{1}{2}(p_{T,jet1} + p_{T,jet2})$

ESV Definitions : $\tau_\perp$ & $Y_{2,3}$

- **Complement of transverse thrust ($\tau_\perp$)** – measure of how highly collimated the momenta in an event along one particular axis –

  $\boxed{\tau_\perp \equiv 1 - T_\perp}$

  Where, thrust in the transverse plane –

  $$T_\perp = \max_{\hat{n}_T}\left(\frac{\sum_i |\vec{p}_{T,i}\cdot\hat{n}_T|}{\sum_i |\vec{p}_{T,i}|}\right),\ 0 \geq \tau_\perp \geq 1 - 2/\pi$$

- $\vec{p}_{T,i} \rightarrow$ transverse momentum of the i^{th} final state object
- $\hat{n}_T \rightarrow$ thrust direction

- **Third-Jet Resolution parameter ($Y_{2,3}$)** –

  $$\boxed{Y_{2,3} = \frac{\max\left(p^2_{T,3}, \left\{\min(p_{T,i}, p_{T,j})^2 \cdot \frac{(\Delta R_{ij})^2}{R^2}\right\}\right)}{P^2_{12}}}$$

- $(\Delta R_{ij})^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$
- $p_{T,3} \rightarrow$ transverse momentum of the third object
- $P_{12} \rightarrow$ scalar sum of the transverse momentum of the two remaining object

ESV Definitions : B_T, ρ_{Tot} & ρ^T_{Tot}

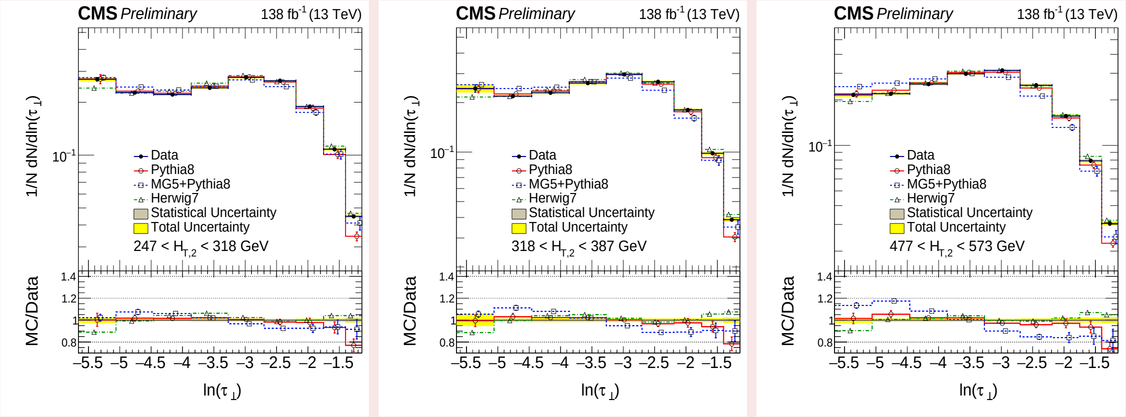
- **Total Jet Broadening** (B_T) – measure of the fraction of energy perpendicular to the thrust axis –
  $\boxed{B_T \equiv B_U + B_L}$, where
  $B_X = \frac{1}{2P_T} \sum_{i \in C_X} p_{T,i} \sqrt{(\eta_i - \eta_X)^2 + (\phi_i - \phi_X)^2}$,
  $\eta_X = \frac{\sum_{i \in C_X} p_{T,i}\, \eta_i}{\sum_{i \in C_X} p_{T,i}}$, $\quad \phi_X = \frac{\sum_{i \in C_X} p_{T,i}\, \phi_i}{\sum_{i \in C_X} p_{T,i}}$

- Each event can be divided into two parts –
  Upper part (U): $\vec{p_T} \cdot \vec{n_T} > 0$
  Lower part (L): $\vec{p_T} \cdot \vec{n_T} < 0$
- $P_T \rightarrow$ scalar sum of the transverse momentum of all the input objects

- **Total Jet Mass** (ρ_{Tot}) –
  $\boxed{\rho_{Tot} \equiv \rho_U + \rho_L}$
- **Total Transverse jet mass** (ρ^T_{Tot}) $\rightarrow$ Transverse component of the ρ_{Tot}

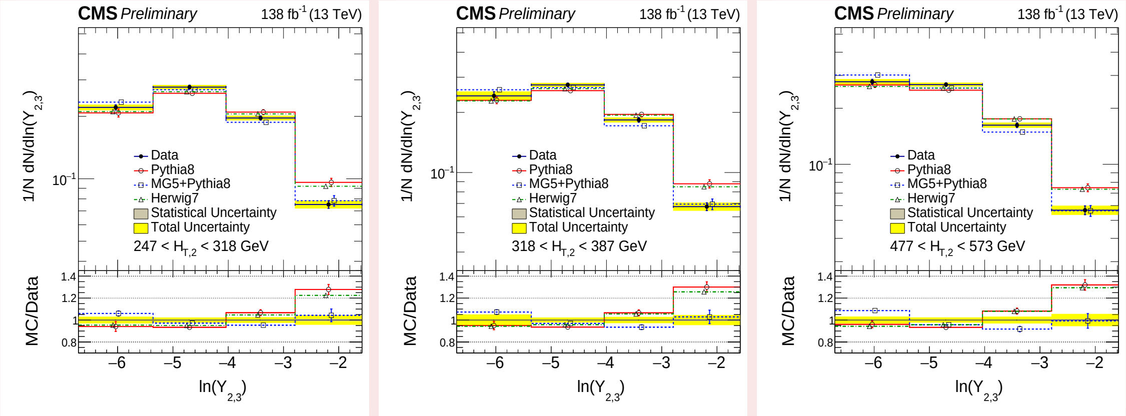
- $\rho_x \equiv \frac{M_x^2}{P^2}$
- $M_x \rightarrow$ Invariant mass of the constitutes of the jets in X (X = U, L)
- $P \rightarrow$ scalar sum of $\vec{p}$ of all constituents

Results : $\tau_{\perp}$



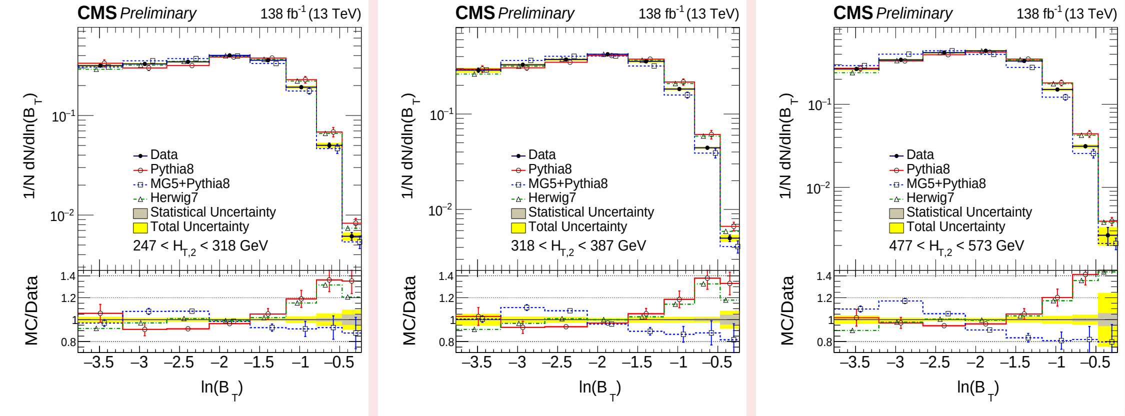
- **Pythia8** – **Good agreement** with data for both spherical and back-to-back jets regions
- **Herwig7** – **Overestimates** the multi-jet, spherical events
- **MG5+Pythia8** – **Good agreement** for lower $H_{T,2}$, **underestimates** the multi-jet, spherical events, for higher $H_{T,2}$

Results : $Y_{2,3}$



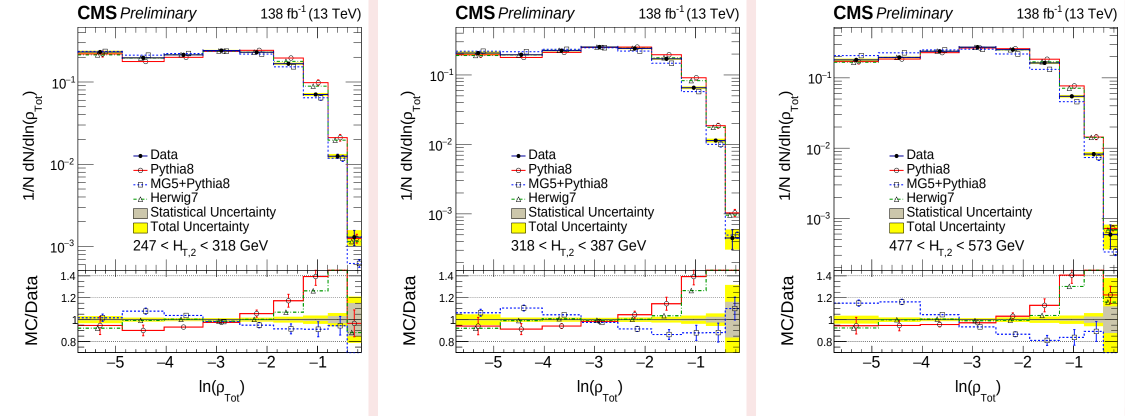
- **PYTHIA8** – **Overestimates** the multi-jet regions
- **HERWIG7** – **Overestimates** the multi-jet regions
- **MG5+PYTHIA8** – **Smaller disagreement**, reduces for higher $H_{T,2}$

Results : B_T



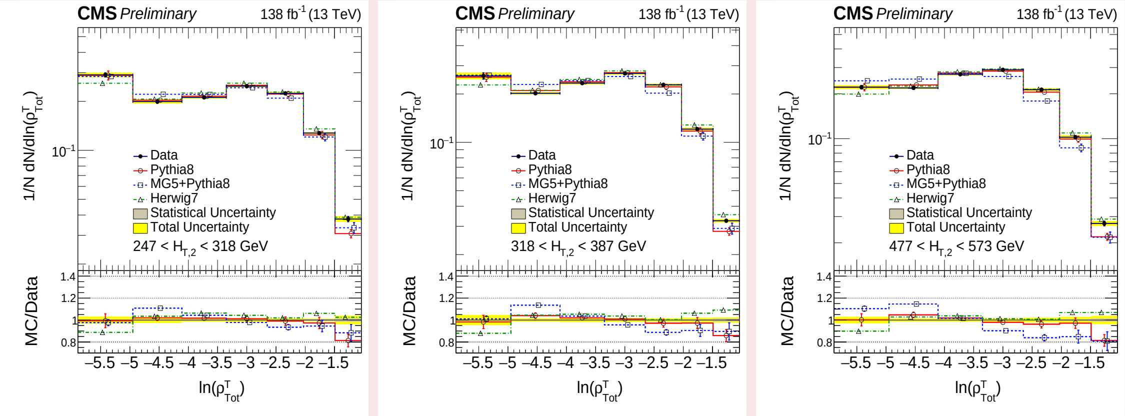
- **Pythia8** – **Overestimates** the multi-jet regions
- **Herwig7** – **Overestimates** the multi-jet regions
- **MG5+Pythia8** – **Good agreement** for the low $H_{T,2}$, **underestimates** spherical events for higher $H_{T,2}$

Results : ρ_{Tot}



- **Pythia8** – **Overestimates** the spherical events
- **Herwig7** – **Overestimates** the spherical events
- **MG5+Pythia8** – **Underestimates** back-to-back two-jet events for lower $H_{T,2}$ ranges but **overestimates** for higher $H_{T,2}$

Results : ρ^{T}_{Tot}



- **Pythia8** – **Good agreement** with data
- **Herwig7** – **Slightly overestimates** the multi-jet region
- **MG5+Pythia8** – **Underestimates** spherical events, **agreement deteriorates** for higher $H_{T,2}$

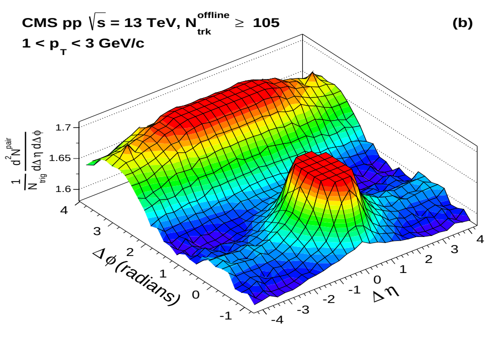
Summary : SMP-22-004

- **Unfolded results** are compared with MC predictions → **Pythia8**, **Herwig7**, **MadGraph5_aMC@NLO**
- **Total uncertainty** → 5-10%
- Major source of uncertainty → **JES** and **MC Models**
- **Pythia8** : **Overall good agreement** with data – $\tau_{\perp}$ and ρ^{T}_{Tot}. **Slightly overestimates** the multijet region for Y_{23}, ρ_{Tot} and B_T
- **Herwig7** : **Overall good agreement** with data for the all ESVs. Agreements are similar to **Pythia8**
- **MadGraph5_aMC@NLO** : **Overall good Agreement** with data, **slightly better** compared to **Pythia8** & **Herwig7**. **Disagreements** observed for multijet events for Y_{23}, ρ_{Tot} and B_T opposite compared to **Pythia8** & **Herwig7**
- **Unfolded results** can be used for **MC generator tuning**

Measurement of event shapes in minimum bias events from pp collisions at 13 TeV

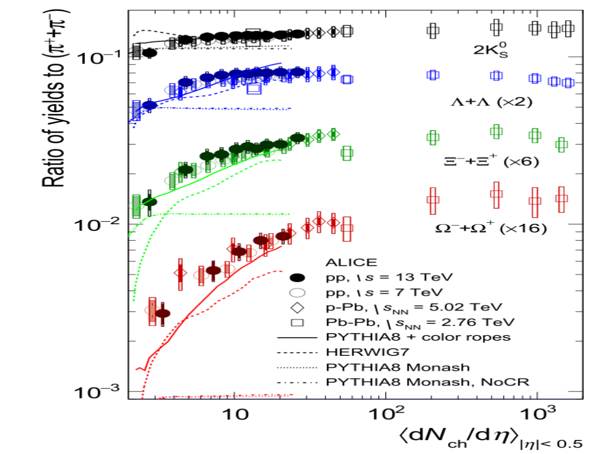
CMS-PAS SMP-23-008

Introduction & Motivation



Phys. Rev. Lett. 116 (2016) 172302

- Unexpected **particle production** across η, with $\Delta\phi \sim 0 \rightarrow$ in **high multiplicity jets**
- Mismodelling of **strangeness production** in pp collisions
- Increase in **strange particle** as a function of **particle multiplicity** $\rightarrow$ not predicted by MC



Eur. Phys. J. C 80, 693 (2020)

- Affects the **detector response** $\rightarrow$ effects **ESVs**
- Indicates mismodelling in **hadronization** & potential **quark-gluon** plasma effects

Introduction & Motivation

Event shapes as functions of charged particle momentum

- Event shape observables:
  - Variables describing the **"shape"** of the events
    $\rightarrow$ Functions of the **momentum of the final state particles**
- Eight ESVs are calculated – ***Charged particle multiplicity (N)***, ***Invariant mass of charged particles ($\sqrt{s}$)***, ***Sphericity (S)***, ***Transverse Sphericity (S_T)***, ***Broadening (B)***, ***Thrust (T)***, ***Transverse Thrust (T_T)***, ***Isotropy***
- **Unfolded results** are compared with **Pythia8**, **Herwig7**, **EPOS-LHC** MC models

Data & Event Selection

- **Data** : **2018 low pileup run** $\rightarrow \mathcal{L} \sim 64\ \mu b^{-1}$
- **Charged particles** $\rightarrow N_{trk} \geq \mathbf{2}$, $p_T > \mathbf{0.5\ GeV}$ & $|\eta| < \mathbf{2.4}$
- No clustering

ESV Definitions : S, S_T & B

- **Sphericity** ($S^{\alpha\beta}$) – measure of how isotropically the momenta are distributed
  $$\boxed{S^{\alpha\beta} = \frac{\sum_i p_i^\alpha p_i^\beta}{\sum_i |\vec{p}_i|^2}}$$
  Sphericity constructed from two smallest eigenvalues λ_2 & λ_3; $S = \frac{3}{2}(\lambda_2 + \lambda_3)$
- **Transverse Sphericity** (S_T)– Sphericity in the transverse plane

- $\alpha, \beta \in x, y, z \rightarrow$ cartesian coordinates

- **Broadening** (B) – measure of the fraction of energy perpendicular to the thrust axis –
  $\boxed{B \equiv B_L + B_R}$, where
  $$B_L = \sum_{i \in L} \frac{|\vec{p_i} \times \vec{n}|}{\sum_i |\vec{p_i}|}, \; B_R = \sum_{i \in R} \frac{|\vec{p_i} \times \vec{n}|}{\sum_i |\vec{p_i}|}$$

- Thrust axis defines left(L) and right(R) hemisphere of the event

ESV Definitions : T, T_T & Isotropy

- **Thrust** (T) – measure of how highly collimated the momenta in an event along one particular axis

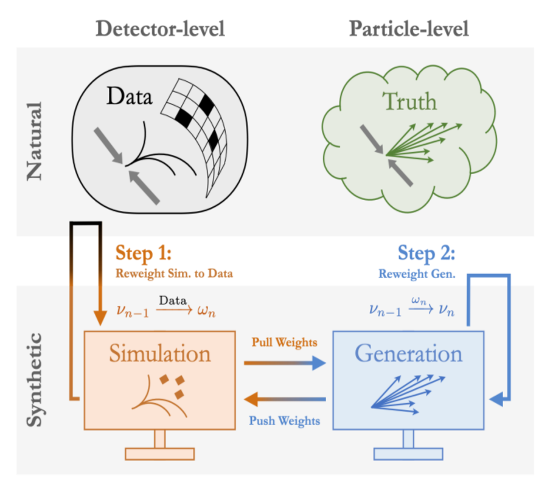
$$T = 1 - \max_{\hat{n}} \left(\frac{\sum_i |\vec{p}_i \cdot \hat{n}|}{\sum_i |\vec{p}_i|} \right)$$

- **Transverse Thrust** (T_T) – Thrust in the transverse plane

- $\vec{p}_i \rightarrow$ momentum of the i^{th} final state object
- $\hat{n} \rightarrow$ thrust direction

- **Isotropy** – Measure of the how isotropically energy is distributed in an event

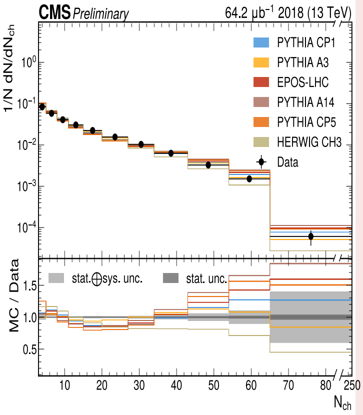
ML-based Unfolding

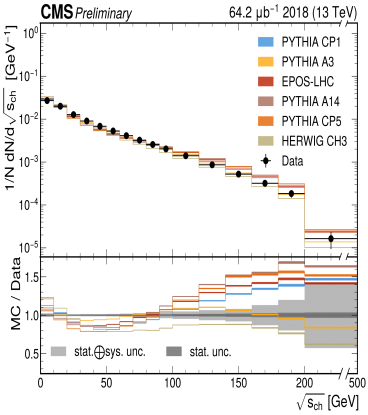


arXiv:1911.09107
arXiv:2105.04448

- **Unfolding** → **Machine Learning** based
- **Multi fold:**
  - **Input** → values of 8 observables for every event in simulation and data
  - **Output** → reweighted simulated events approximating data
  - Results are unbinned weighted events
- **Two steps of unbinned reweighting:**
  - Weight MC → data at detector level
  - Weight original MC → reweighted MC at generator level
  - Extra 2 steps added → deal with the selection efficiency and signal acceptance → Repeat in iterations

Results : N_{ch} & $\sqrt{s_{ch}}$

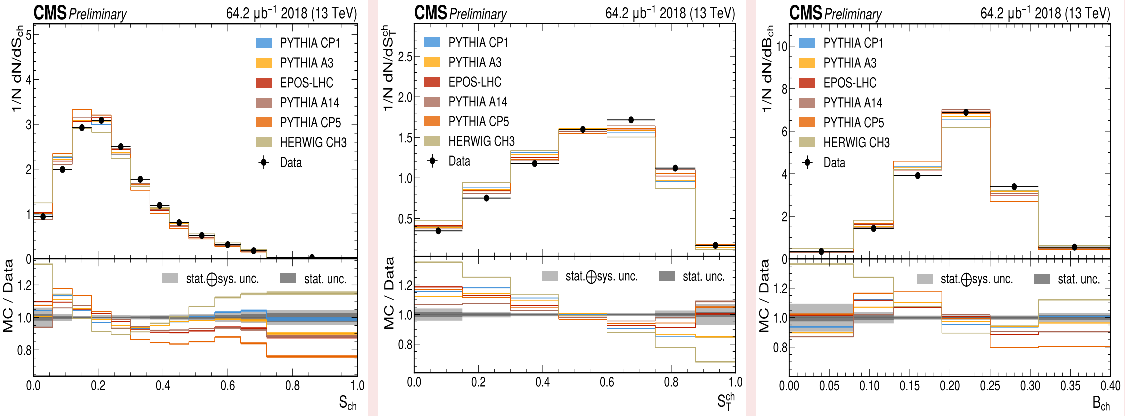




Uncertainty source	Relative uncertainty
Bias from regularization	1.2%
Mismodelling of the migration	1.6%
Track reconstruction efficiency	1.7%
MC statistics	0.26%
Data statistics	0.26%

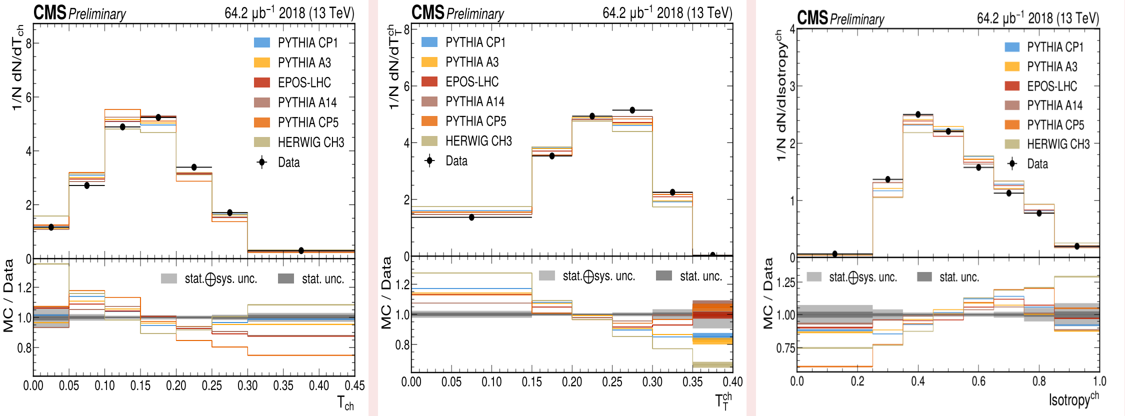
- **Unfolded** N_{ch} & $\sqrt{s_{ch}}$ are compared with different **MC predictions**
- Data is **more accumulated** in the **middle** N_{ch} **region** ($\sim$ 20)

Results : S_{ch}, S_T^{ch} & B_{ch}



- **Unfolded** S_{ch}, S_T^{ch} **&** B_{ch} are compared with different **MC predictions**
- **Mismodelling** observed for all ESVs

Results : T_{ch}, T_T^{ch} & $Isotropy^{ch}$



- Data-MC **discrepancy is largest** in the **middle** N_{ch} region
- For N_{ch} $\lesssim$**10** and **low** $\sqrt{S_{ch}}$ $\rightarrow$ all generators **overestimate** the fraction of events
- For **high** N_{ch} and **high** $\sqrt{S_{ch}}$ $\rightarrow$ predictions of the MC generators **diverge**

Summary : SMP-23-008

- Minimum bias collisions may be the home of **interesting physics at the LHC**
- **Eight different ESVs** are **unfolded** and compared with **MC predictions**
- **Mismodelling** is observed for **all ESVs**
- **Unfolded results** can be used for **further investigations**

Conclusions

- **CMS** provided many measurements to **understand QCD using Run-2 data** – **event and jet shapes**, **charged particle production**, ...
- The measurements focus on different aspects of particle production in **hadron (and nuclear) collisions**
- Presented two recent **event shape measurements** –
  - ***Measurement of event shape variables using charged particles inside jets in proton-proton collisions at*** $\sqrt{s} = 13$ ***TeV***
  - ***Measurement of event shapes in minimum bias events from pp collisions at 13 TeV***
- **Unfolded results** can be used for further **theoretical predictions & generator tuning**
- Study with **Run-3 data** in its way
- Many more exciting results are **coming soon**