Jet substructure studies with ATLAS and CMS

Luka SELEM
On behalf of the ATLAS and CMS collaborations

QCD@LHC conference
Freiburg im Brisgau
07/10/2024

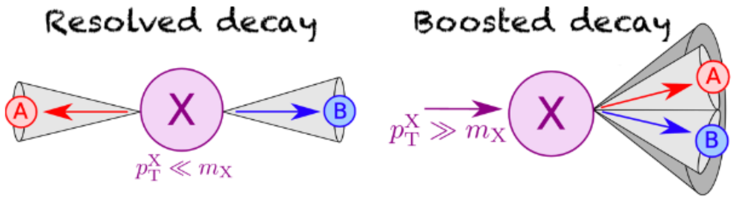
ATLAS EXPERIMENT

CMS Compact Muon Solenoid

Why study jet substructure

Jet substructures to resolve boosted signatures

- From **jet mass** to more complex variables
- **Jet tagging** : Quark/gluon, W/Z, top, higgs
- **Discriminate exotics** models



Jet substructure to test and **improve Monte Carlo prediction**

- Parton shower, hadronisation models
- Improve MC parameter tuning

Study **non-perturbative** QCD and **jet fragmentation**

➔ To be compared to heavy ions QGP

New results to be presented

The Lund jet plane

- for inclusive QCD jets: JHEP 05 (2024) 116 (CMS)
- for top and W jets: arXiv:2407.10879 (ATLAS)
- as a tool to reweight multi-prong jets: CMS-DP-2023-046(CMS)

Improving Monte Carlo prediction

- Top jet substructure in ttbar: Phys. Rev. D 109 (2024) 112016 (ATLAS)
- Lund Subjet Multiplicity: arXiv:2402.13052 (ATLAS)
- Unbinned substructure variables unfolding with OMNIFOLD: arXiv:2405.20041 (ATLAS)

Study of jet fragmentation

- Track function = integrate hadrons correlations: ATLAS-CONF-2024-012 (ATLAS)
- Dead cone effect in b-jets and D^0 jets: CMS-PAS-HIN-24-005, CMS-PAS-HIN-24-007 (CMS)

The Lund Jet Plane

CMS result for inclusive jet selection in Run 2

JHEP 05 (2024) 116

*Measurement of the **primary Lund jet plane density** in proton-proton collisions at $\sqrt{s}$ = 13 TeV*

ATLAS result in ttbar events from Run 2

arXiv:2407.10879

*Measurement of the Lund jet plane in **hadronic decays of top quarks and W bosons** with the ATLAS detector*

CMS Performance Note based on Run 2 data

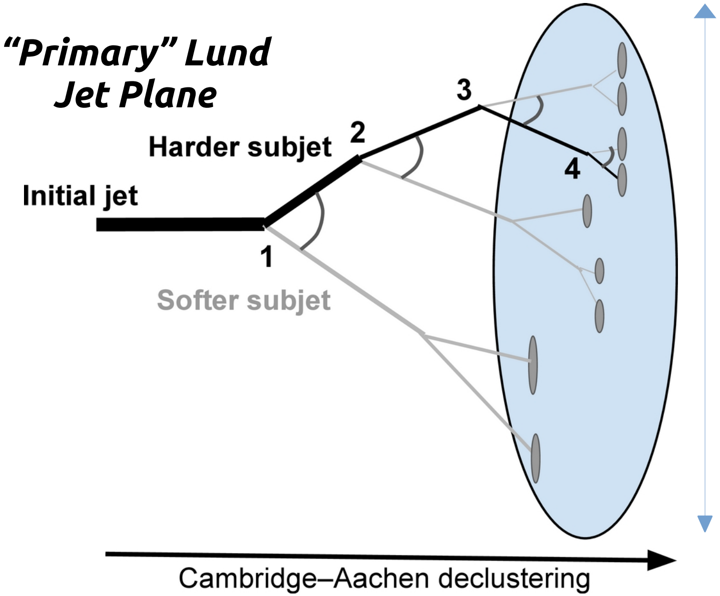
CMS-DP-2023-046

*Lund Plane Reweighting for **Jet Substructure Correction***

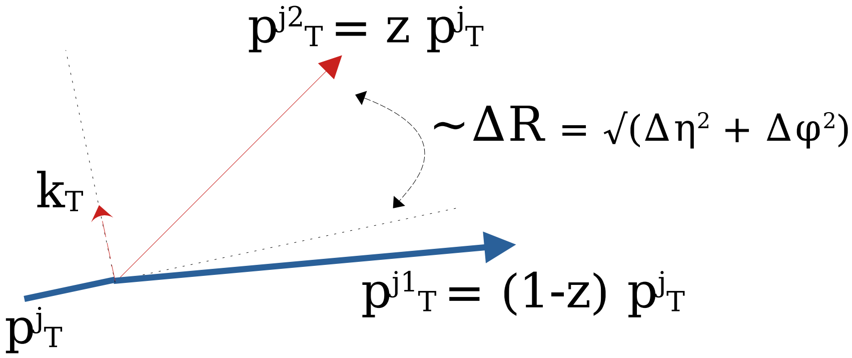
Building the Lund Jet Plane

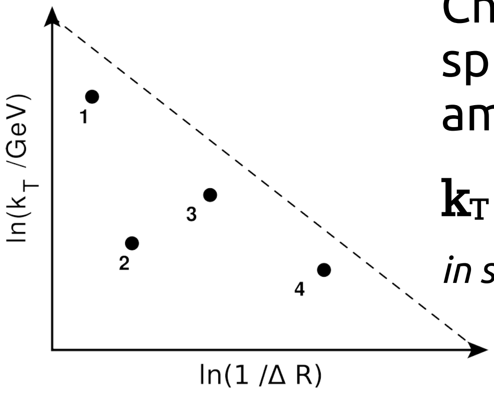
Give an image of a jet and its QCD radiation structure

➔ Represent the successive QCD emissions in the formation of a jet



*Jet constituents clustered with Anti-k_T algorithm with free parameter **R***





Choice of pair of splitting variables among **k_T, z, ΔR** with

$$k_T \sim p_{j1}\, \Delta R = z\, p_j\, \Delta R$$

in soft (z<<1) collinear (ΔR<<1) limit

Angle ordered algorithm based on

$$d_{ij} = \frac{\Delta R_{ij}^2}{R^2}$$

NOTE:** declustering done only using **charged particles

→ use tracks with better angular precision

Properties of the Lund Jet Plane

Subsequent emissions move left to right

➔ Cambridge-Aachen angular ordering

Define emission average density

$$\rho(k_{\mathrm{T}}, \Delta R) \equiv \frac{1}{N_{\mathrm{jets}}} \frac{\mathrm{d}^2 N_{\mathrm{emissions}}}{\mathrm{d}\ln(k_{\mathrm{T}}/\mathrm{GeV})\,\mathrm{d}\ln(R/\Delta R)}$$

Get a per-jet picture

Soft collinear region

➔plateau evolving with running of α_s

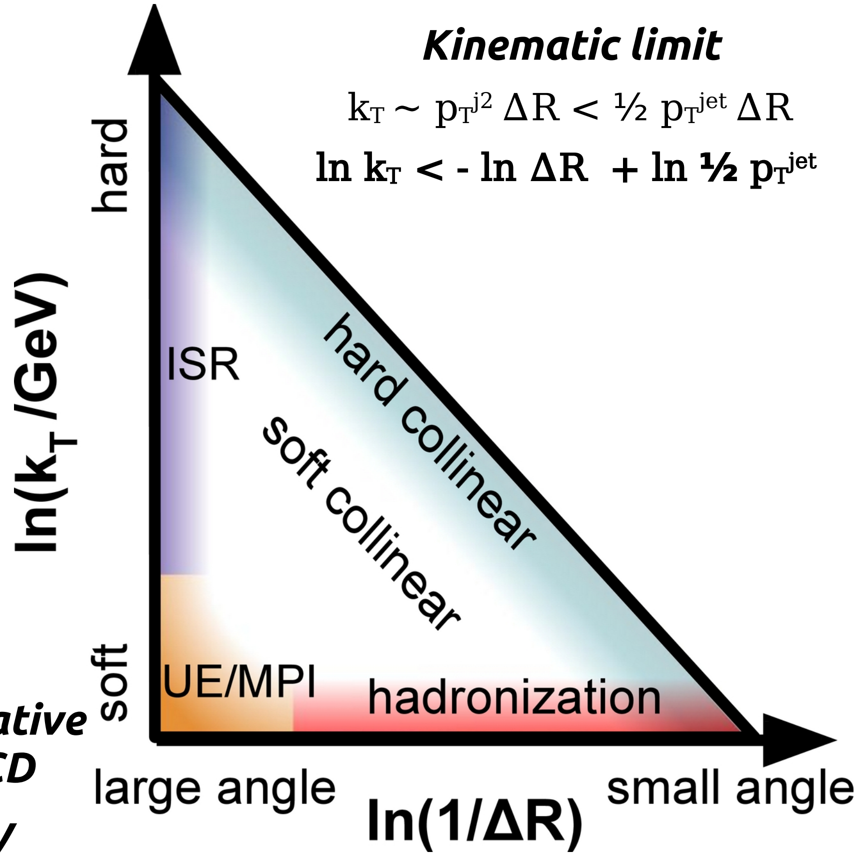
$$\rho(k_{\mathrm{T}}, \Delta R) \approx \frac{2}{\pi} C_{\mathrm{R}} \alpha_{\mathrm{S}}(k_{\mathrm{T}})$$

CR : color factor

Kinematic limit

$k_T \sim p_T^{j2}\, \Delta R < \frac{1}{2}\, p_T^{jet}\, \Delta R$

$\ln k_T < -\ln \Delta R + \ln \frac{1}{2} p_T^{jet}$

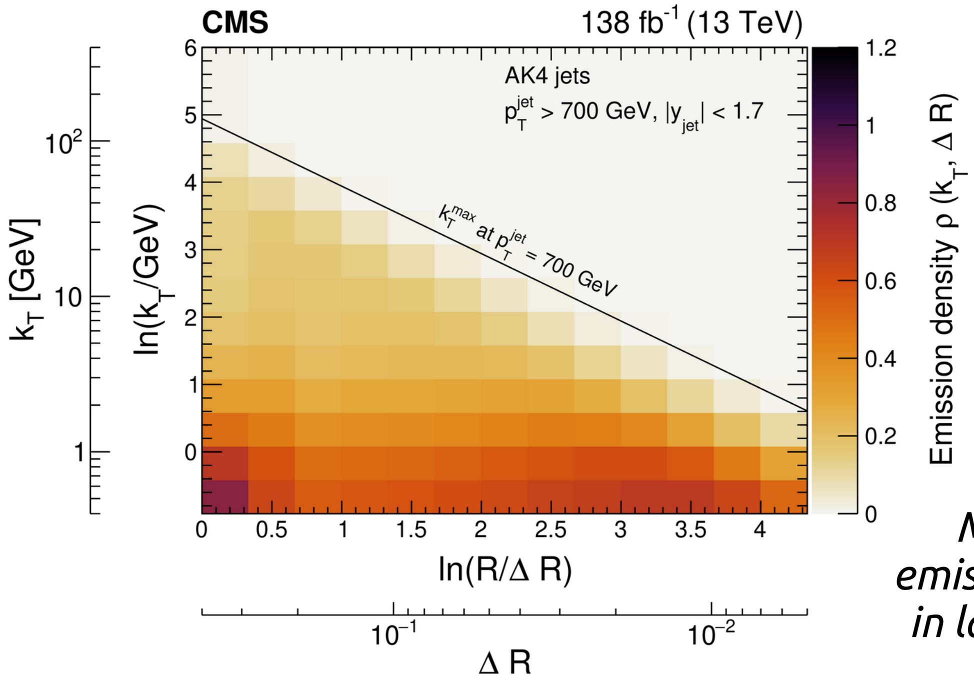


Non-perturbative limit of QCD

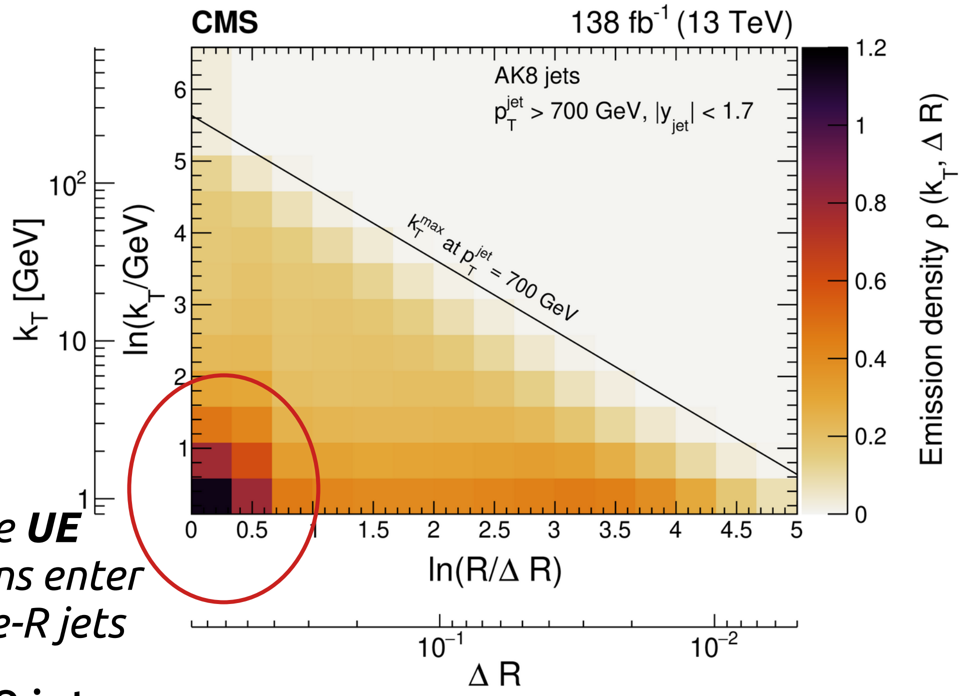
$k_T \sim 1$ GeV

Unfolded Lund Jet Plane

Iterative Bayesian Unfolding for inclusive QCD jets



*More **UE** emissions enter in large-R jets*



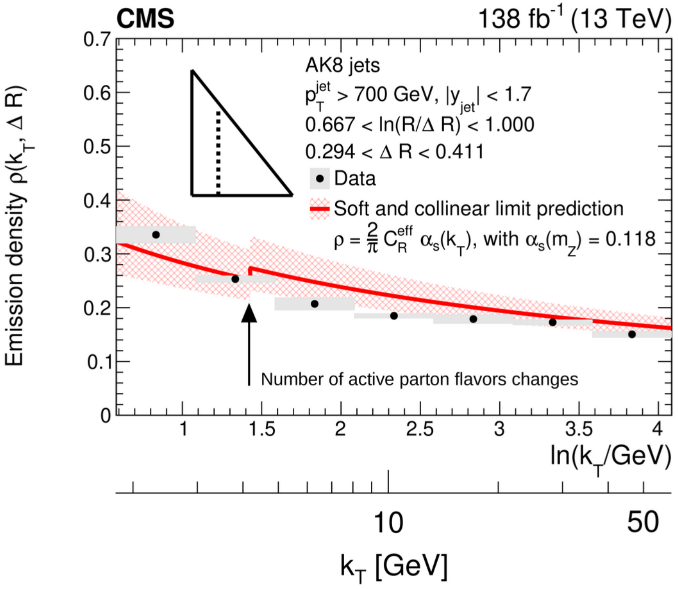
Soft collinear plateau more visible with R=0.8 jets

➔ More phase space available

$$\rho(k_{\mathrm{T}}, \Delta R) \approx \frac{2}{\pi} C_{\mathrm{R}} \alpha_{\mathrm{S}}(k_{\mathrm{T}})$$

***α_s** increases when **k_T** decrease*
(until non-perturbative effects enter)

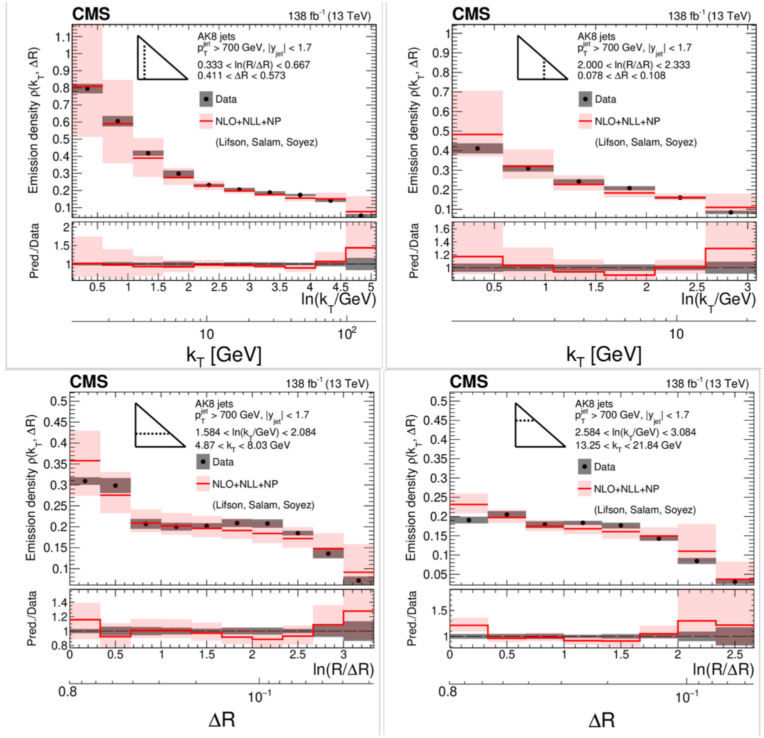
Theory prediction of LJP



Soft collinear one-loop analytical prediction

➔ Running of α_s well modelled

$$\rho(k_{\mathrm{T}}, \Delta R) \approx \frac{2}{\pi} C_{\mathrm{R}} \alpha_{\mathrm{S}}(k_{\mathrm{T}})$$

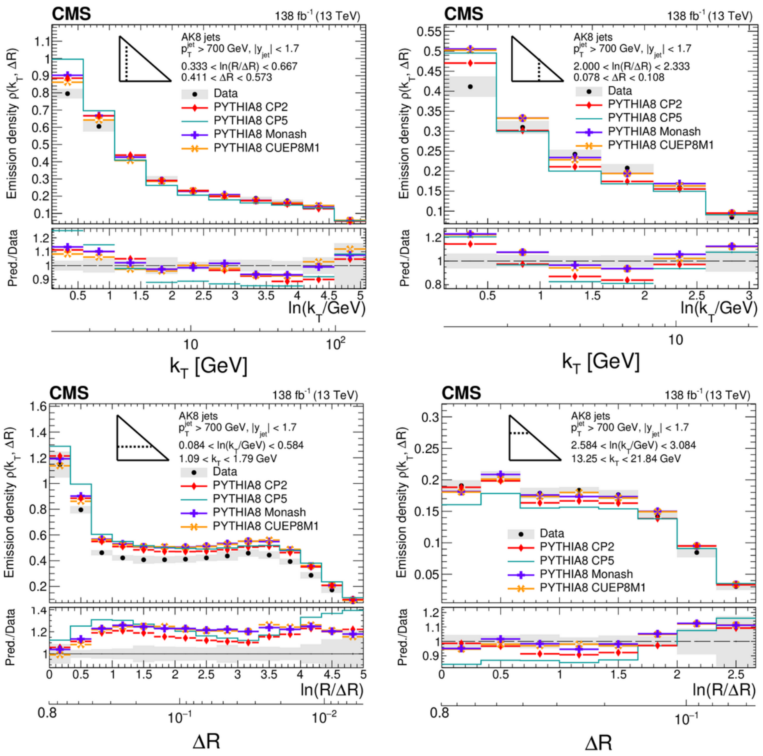


Agreement with NLO + NLL + Non-Perturbative predictions

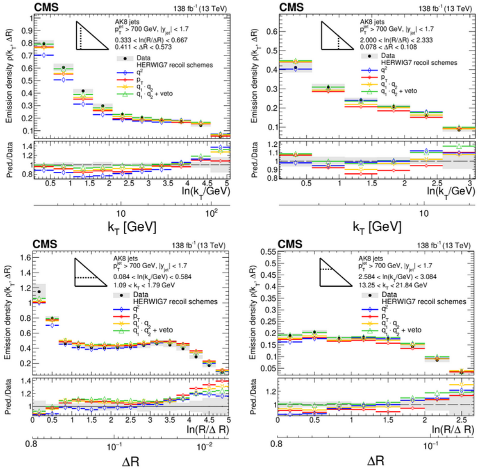
➔ **First principle understanding of LJP**

Benchmark choice for Monte-Carlo

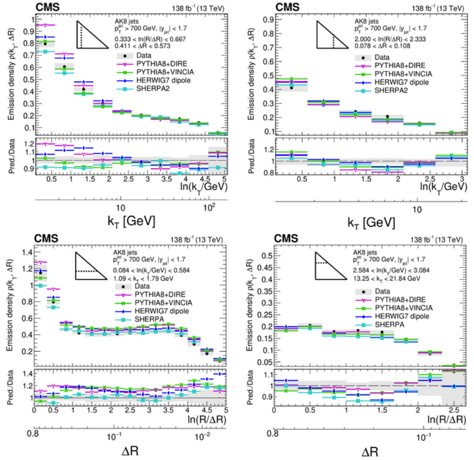
Pythia tuning



Herwig recoil scheme



Dipole shower implementation



CP5 best for multijet, found worst in LJP (jet substructure)

➔ **Complementarity !**

The Lund Jet Plane

CMS result for inclusive jet selection in Run 2

JHEP 05 (2024) 116

*Measurement of the **primary Lund jet plane density** in proton-proton collisions at √s = 13 TeV*

ATLAS result in ttbar events from Run 2

arXiv:2407.10879

*Measurement of the Lund jet plane in **hadronic decays of top quarks and W bosons** with the ATLAS detector*

CMS Performance Note based on Run 2 data

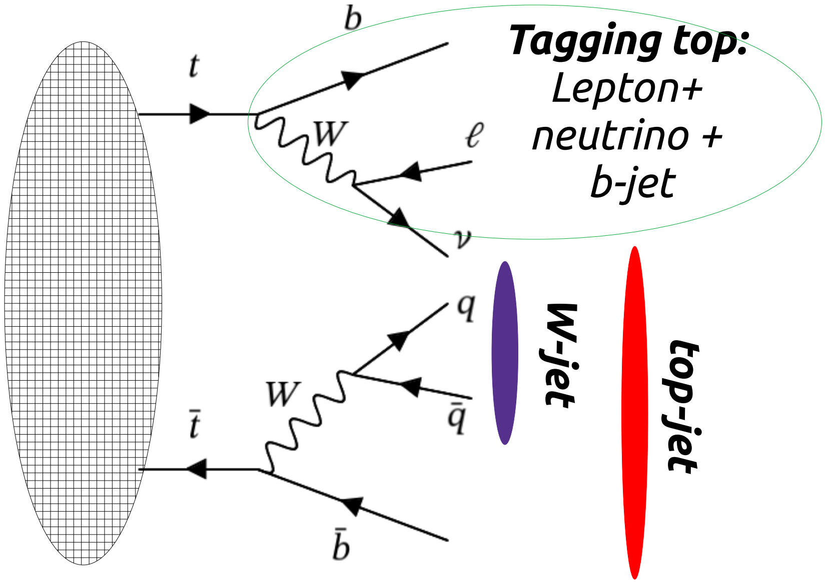
CMS-DP-2023-046

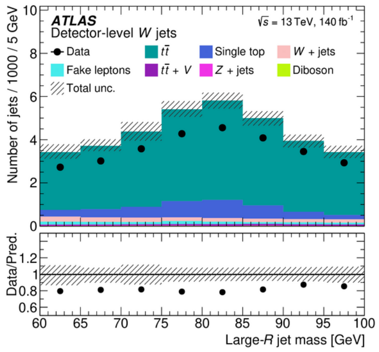
*Lund Plane Reweighting for **Jet Substructure Correction***

Selection of t tbar events

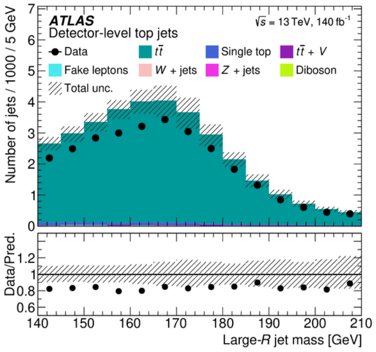
Jet definition	
Constituents	Weighted calorimeter topo clusters
Clustering	Anti-k_T $R = 1.0$
Grooming	Trimming $R_{\text{trim}} = 0.2,\ f_{\text{trim}} = 0.05$
Event selection	
Trigger	Single lepton $(e^{\pm}, \mu^{\pm})$ trigger
W Lepton	$= 1$ passing lepton selection
W Neutrino	$E_T^{\text{miss}} > 60\,\text{GeV},\ E_T^{\text{miss}} + m_T^W > 20\,\text{GeV}$
b-tagged jet	≥ 1 and $\Delta R(\ell, j_{b1}) < 1.5$
Large-R Jet	$p_T^{jet} > 350\,\text{GeV},\ \lvert y\rvert < 1.1$ and $\Delta R(\ell, J) > 2.3$
Large-R jet classification	
top-jet	$m_J > 140\,\text{GeV}$, 2^{nd} b-jet with $\Delta R(J, j_{b2}) < 1.0$
W-jet	$60\,\text{GeV} < m_J < 100\,\text{GeV}$
Backgrounds	
Irreducible tW, V+jets, $t\bar{t}Z$, $t\bar{t}W$, $t\bar{t}H$, VV	MC simulations
Fake leptons	Data-driven *Matrix Method*

Per event, leading p_T Large-R jet classified (W or top) and used to reconstruct the Lund Jet Plane





W-jet mass



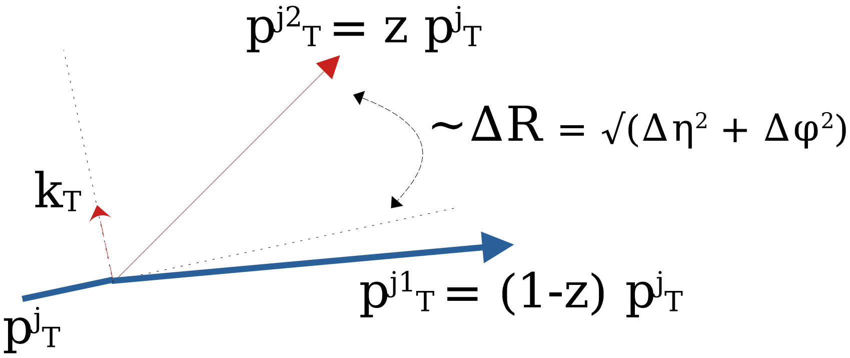
top-jet mass

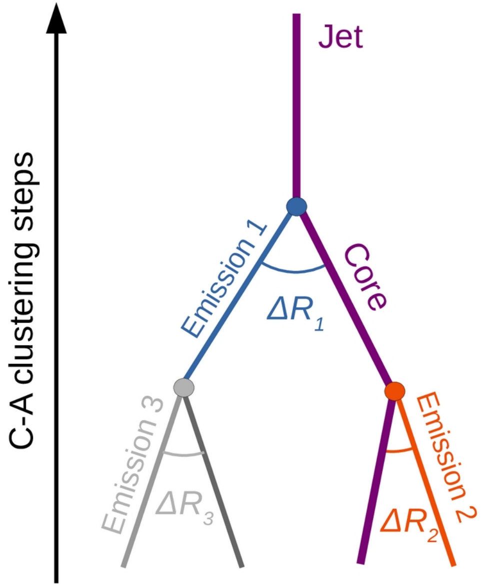
Change of Lund Jet Plane definition

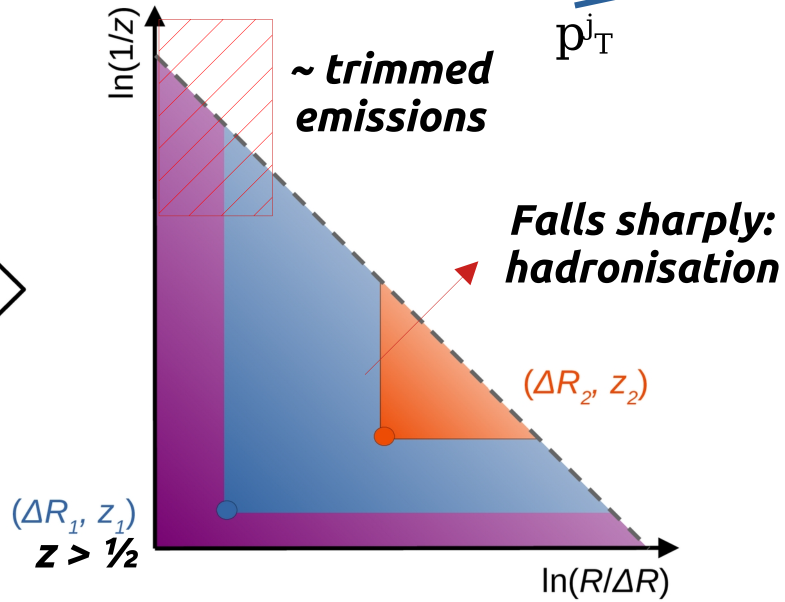
Still Primary, still only charged particles

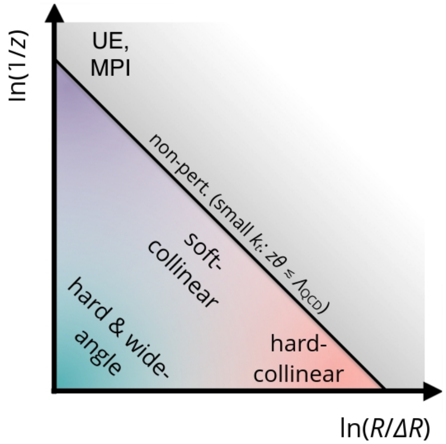
Replacing $\mathbf{k_T}$ by $\mathbf{z}$

➔ $\mathbf{z}$ linked to jet trimming algorithm







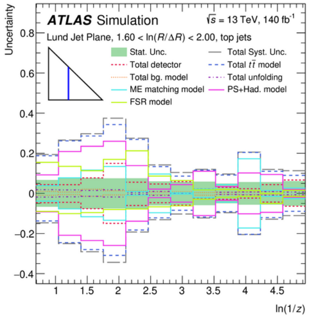


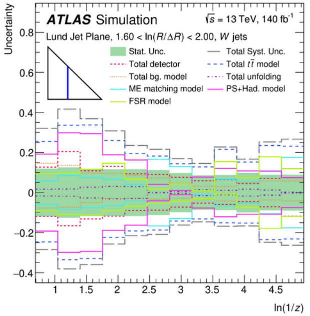
Large bins correlations across the plane

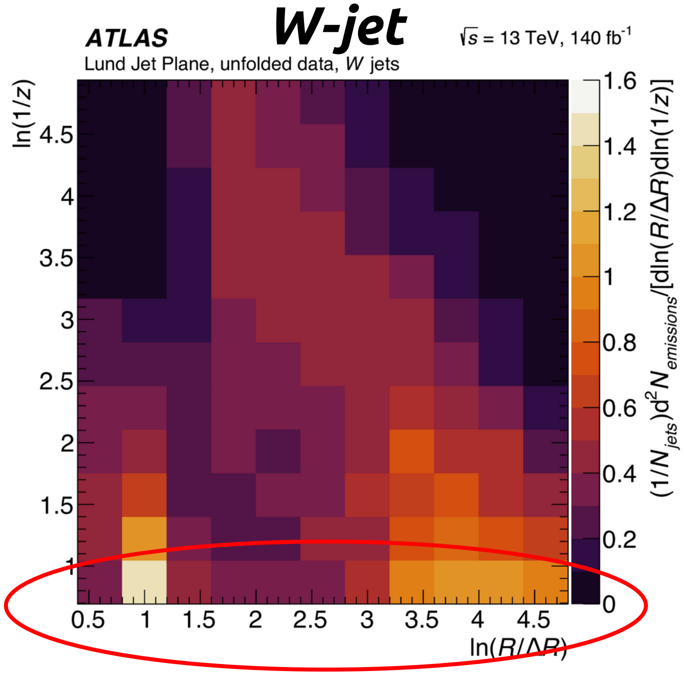
Unfolded Lund Jet Plane

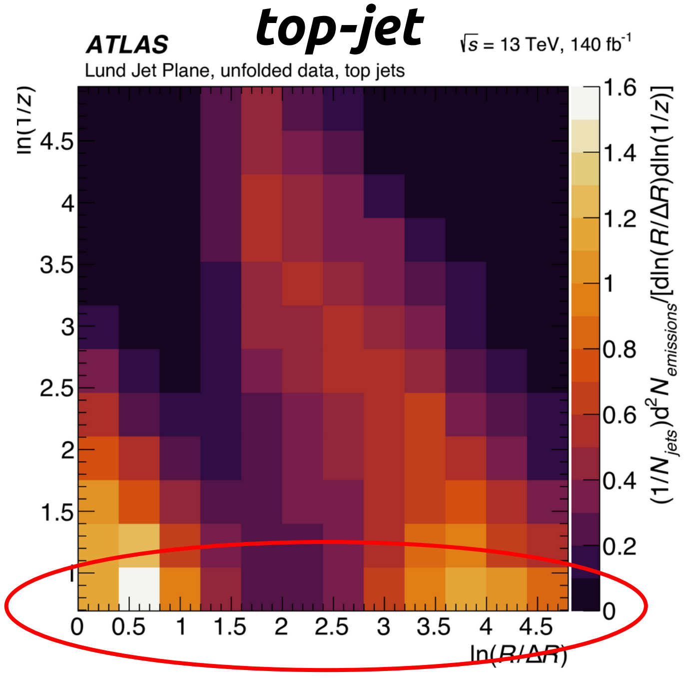
Iterative Bayesian Unfolding to particle level

Uncertainties dominated by modelling systematics (ttbar and PS modelling) at **~10-40%**









W-jets more collimated hard radiations than top-jet

➔ $M_{top} > M_W$

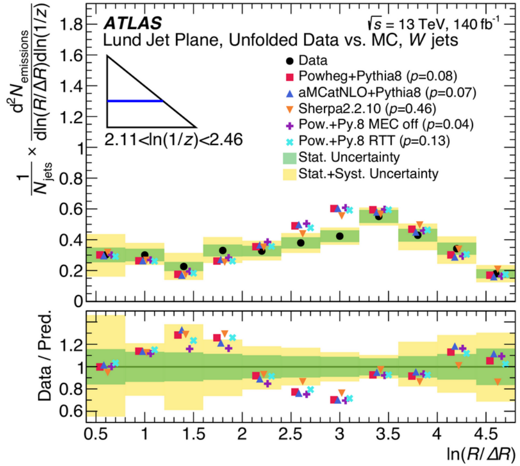
On average **one more emissions** in top-jet than W-jet:

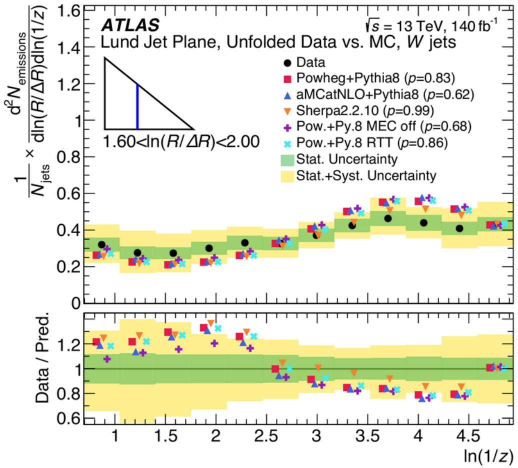
- top-jet : **6.74 ± 0.13**
- W-jet : **6.02 ± 0.22**

Jet tagging ?

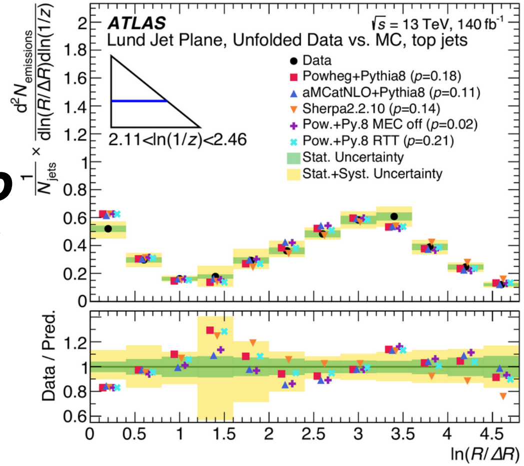
MC benchmarking

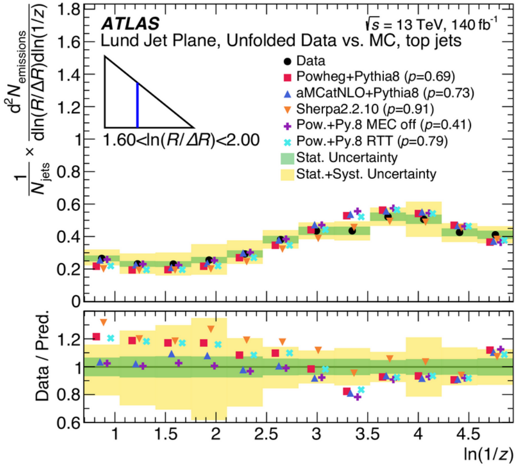
W jet





top jet





χ^2 agreement test between all MC and data across the Lund Jet Plane

- **W-jets : all p-values < 1%**
- Top-jets: all > 1%, best is SHERPA2.2.10 (p-value ~ 33%)

Local p-values in sub-regions are systematically higher

➔ Lost correlations between bins across the plane

Improve tuning of MC for ttbar events

The Lund Jet Plane

CMS result for inclusive jet selection in Run 2

JHEP 05 (2024) 116

*Measurement of the **primary Lund jet plane density** in proton-proton collisions at $\sqrt{s}$ = 13 TeV*

ATLAS result in ttbar events from Run 2

arXiv:2407.10879

*Measurement of the Lund jet plane in **hadronic decays of top quarks and W bosons** with the ATLAS detector*

CMS Performance Note based on Run 2 data

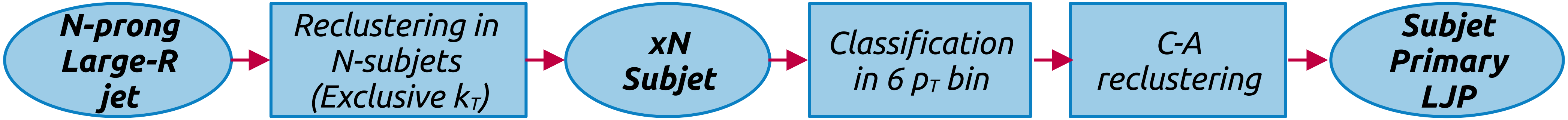
CMS-DP-2023-046

*Lund Plane Reweighting for **Jet Substructure Correction***

Calibrating multi-prong jets

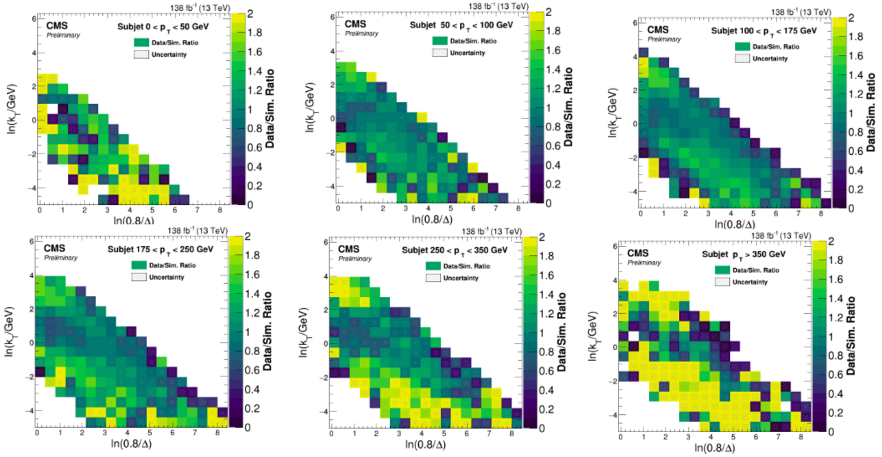
Calibrate multi-prong jets from exotic signal

➔ Reweight Lund Jet Plane of subjets



$$\frac{\textit{Data LJP}}{\textit{Simulated Signal LJP} - \textit{Simulated Background LJP}}$$

Maps in 6 jet p_T bins



Maps made with SM subjets

- Used to reweight each splitting in subjets

➔**Calibration of multi-prong jets from exotic signals**

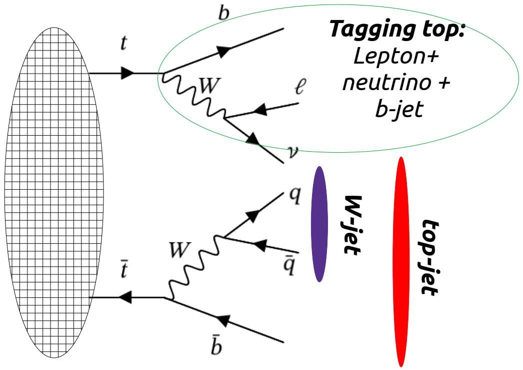
Validation

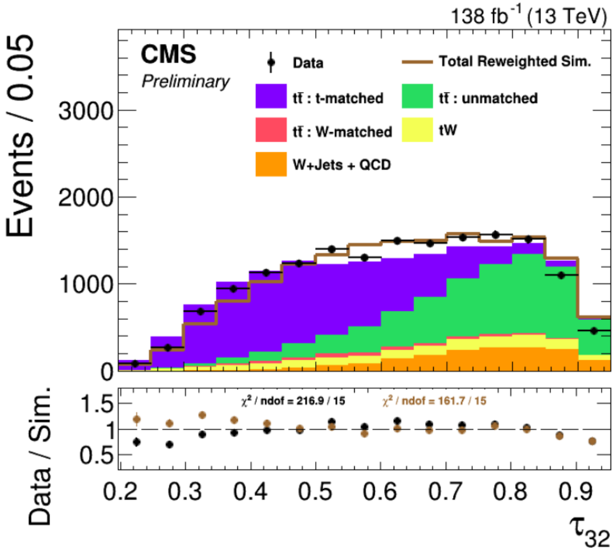
Using semi-leptonic ttbar channel

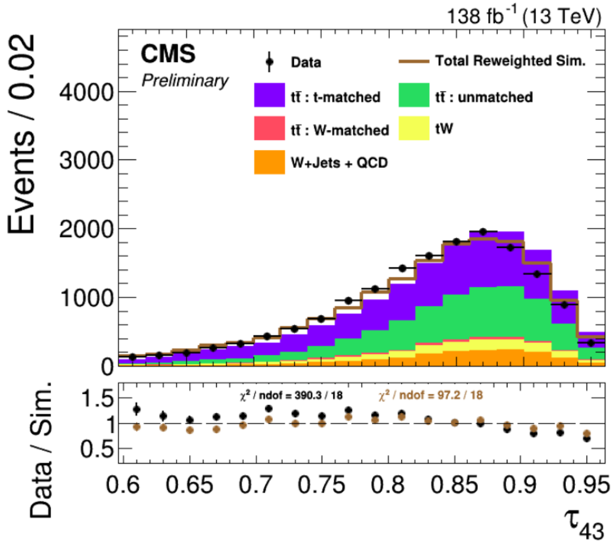
➔ Similar selection as ATLAS analysis

Build reweighting LJP maps with W jets (2-prongs)

➔ Apply to top-jets (3-prong)







Improvement extrapolates to higher number of prongs

➔ Particulary visible in 3-prongs related variable

Main uncertainty from **matching** reco subjet to truth quark

➔ Increases with number of prongs

Improving Monte Carlo predictions

ATLAS result in ttbar events from Run2
Phys. Rev. D 109 (2024) 112016
*Measurement of **jet substructure in boosted tt events** with the ATLAS detector using 140 fb^{-1} of 13 TeV pp collisions*

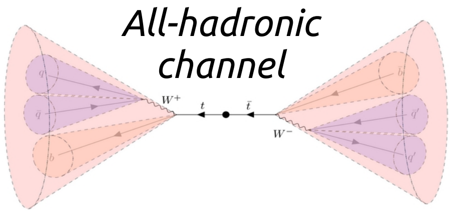
ATLAS result in dijet events from Run 2
arXiv:2402.13052
*Measurements of **Lund subjet multiplicities** in 13 TeV proton-proton collisions with the ATLAS detector*

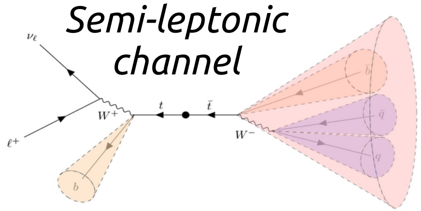
ATLAS result in Z+jets events from Run 2
arXiv:2405.20041
*A **simultaneous unbinned differential cross section** measurement of twenty-four Z+jets kinematic observables with the ATLAS detector*

Studying Large-R jets from top quark decay

Goal: study in **boosted** regime the substructure of top jets in **ttbar events**

➔ Average jet **pT ~ 500GeV**





8 jet substructure variables variables from **“ghost associated” tracks** unfolded with **Iterative Bayesian Unfolding**

(definition in extra material)

- τ_{32}
- τ_{21}
- τ_3
- $ECF2$
- D_2
- C_3
- $p_T^{d,*}$
- LHA

*Most **sensitive** variables in*
W-jet/top-jet multivariate discriminator
used by ATLAS (→ τ_{32}, D_2)

Momentum splitting ~ Sensitive to hadronisation model (→ $p_T^{d,*}$)

Goodness of modelling of variables

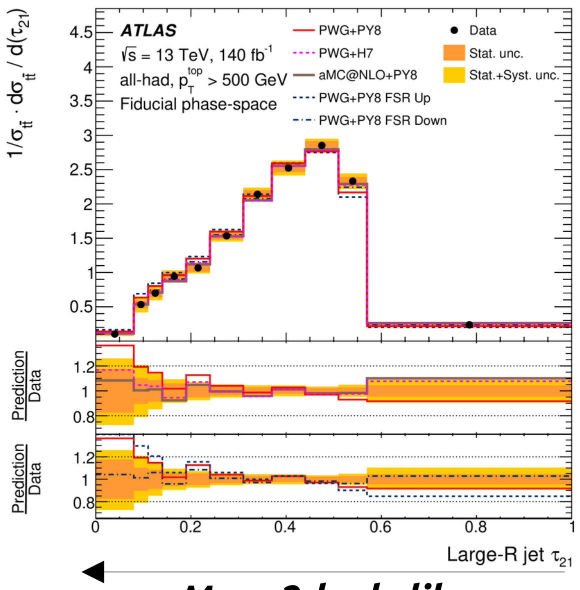
Two-body related variables (**τ_{21}, D_2**) well modelled

Three-body variables (**τ_{32}, τ_3, C_3**) have **MC overestimate** 3-body structure

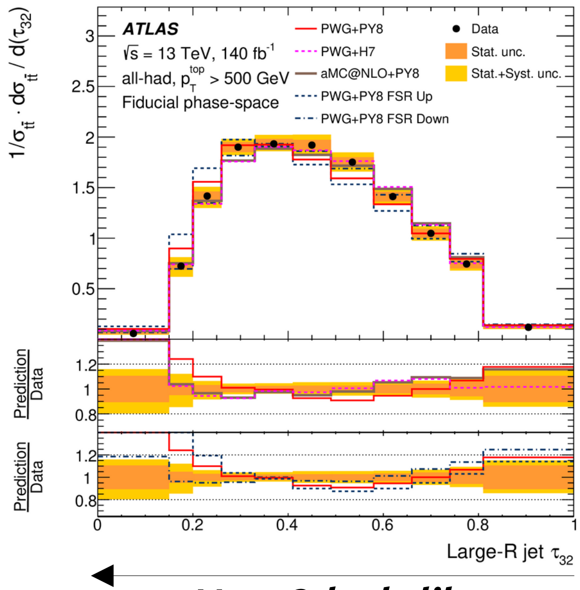
➔ Closer to 0

Poor modelling of the p_T dispersion **p_T^{d*}**

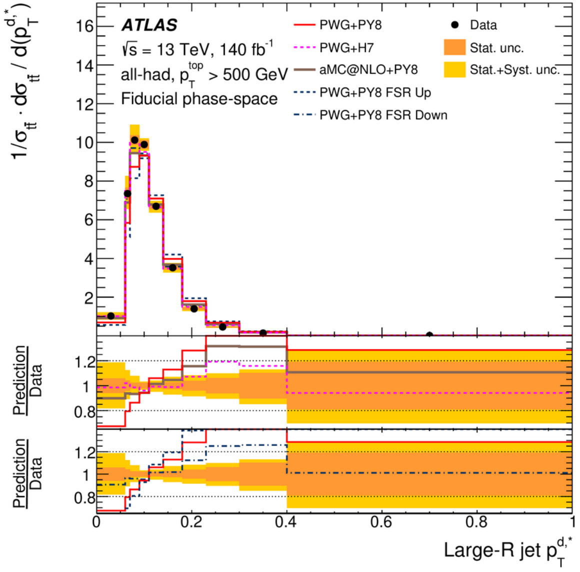
LHA and **ECF2** well modelled



More 2-body like



More 3-body like



Quantitative agreement : χ^2 test

Pythia FSR Up disfavoured,
FSR Down improves agreement
➔ **Data favours increasing α_s^{FSR}**

Better than nominal

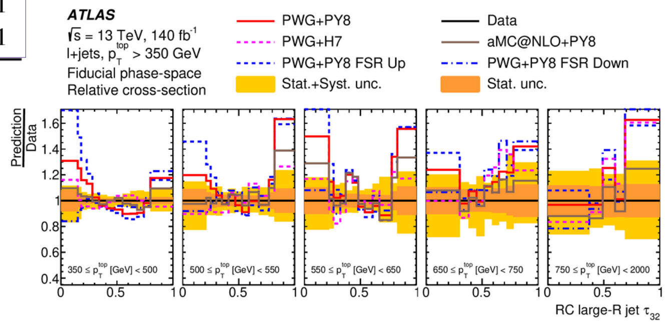
Nominal | *~~Pythia~~ → Herwig* | *~~Powheg~~ → Madgraph* | *Bad Agreement* | *Improved Agreement*

Observable	PWG+PY8		PWG+H7		aMC@NLO+PY8		PWG+PY8(FSR Up)		PWG+PY8(FSR Down)	
	χ^2/NDF	p-value	χ^2/NDF	p-value	χ^2/NDF	p-value	χ^2/NDF	p-value	χ^2/NDF	p-value
τ_{32}	54/12	<0.01	19/12	0.09	15/12	0.24	165/12	<0.01	40/12	<0.01
τ_{21}	14/14	0.41	7/14	0.92	16/14	0.32	42/14	<0.01	8/14	0.91
τ_3	36/11	<0.01	42/11	<0.01	14/11	0.23	130/11	<0.01	23/11	0.02
$ECF2$	25/18	0.13	13/18	0.78	15/18	0.69	31/18	0.03	24/18	0.14
D_2	20/16	0.20	17/16	0.39	20/16	0.20	37/16	<0.01	15/16	0.49
C_3	11/14	0.65	6/14	0.97	3/14	1.00	35/14	<0.01	3/14	1.00
$p_T^{d,*}$	27/12	<0.01	10/12	0.58	11/12	0.53	56/12	<0.01	24/12	0.02
LHA	14/17	0.65	9/17	0.92	20/17	0.29	14/17	0.69	19/17	0.32
D_2 vs. m^{top}	61/42	0.03	62/42	0.02	59/42	0.05	118/42	<0.01	44/42	0.37
D_2 vs. p_T^{top}	71/56	0.08	68/56	0.13	70/56	0.11	107/56	<0.01	93/56	<0.01
τ_{32} vs. m^{top}	153/42	<0.01	72/42	<0.01	56/42	0.07	413/42	<0.01	77/42	<0.01
τ_{32} vs. p_T^{top}	153/50	<0.01	103/50	<0.01	57/50	0.23	360/50	<0.01	114/50	<0.01

3-body mismodelling

Sensitive to hadronisation model

Correlations with m^{top} and p_T^{top}



D_2 and τ_{32} in tension with data **in 2D distributions**
➔ Modelling improvement to **yield lower theoretical uncertainties in top-tagging**

Improving Monte Carlo predictions

ATLAS result in ttbar events from Run2
Phys. Rev. D 109 (2024) 112016
*Measurement of **jet substructure in boosted $t\bar{t}$ events** with the ATLAS detector using 140 fb^{-1} of 13 TeV pp collisions*

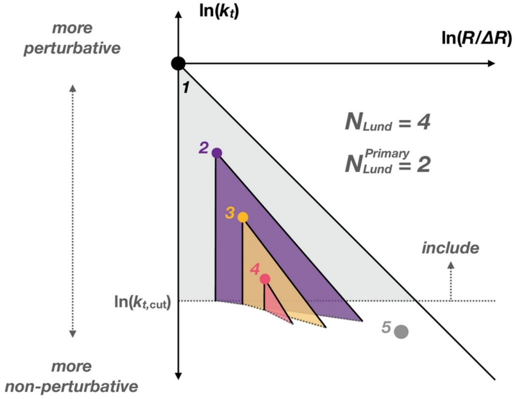
ATLAS result in dijet events from Run 2
arXiv:2402.13052
*Measurements of **Lund subjet multiplicities** in 13 TeV proton-proton collisions with the ATLAS detector*

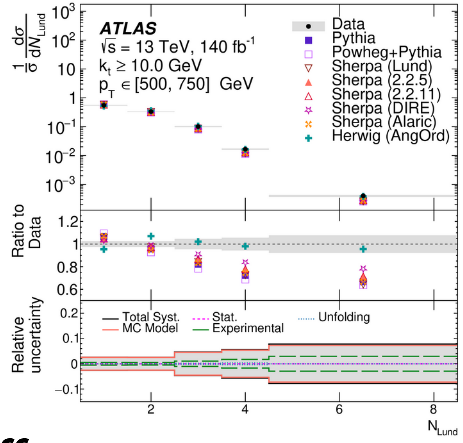
ATLAS result in Z+jets events from Run 2
arXiv:2402.13052
*A **simultaneous unbinned differential cross section** measurement of twenty-four Z+jets kinematic observables with the ATLAS detector*

Lund Subjet multiplicity at various k_t cut-off

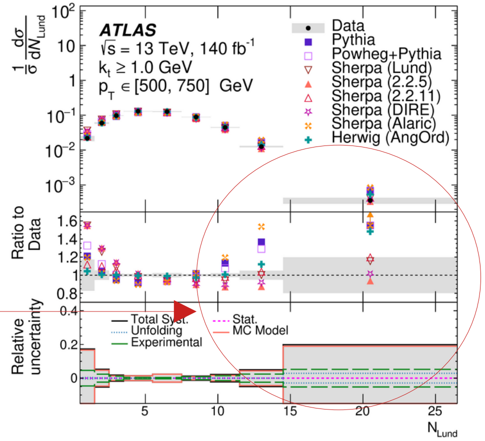
Unfold Lund Subjet Multiplicity
(in dijet events)

➔Per-jet number of emission in Primary plane $N_{Lund}^{Primary}$ and in **all Lund planes N_{Lund}**





k_t cut-off goes down



Varied k_t cut-off : from 0.5 GeV to 100 GeV

➔ Rescaled to account for neutral particles

$$k_t = (p_T^{\text{all}} / p_T^{\text{charged}}) k_t^{\text{charged}}$$

MC benchmark

– **HERWIG Angular Ordered best overall**

– SHERPA (DIRE) better at high multiplicity, low cut-off

➔Better modelling of **non-perturbative** effects

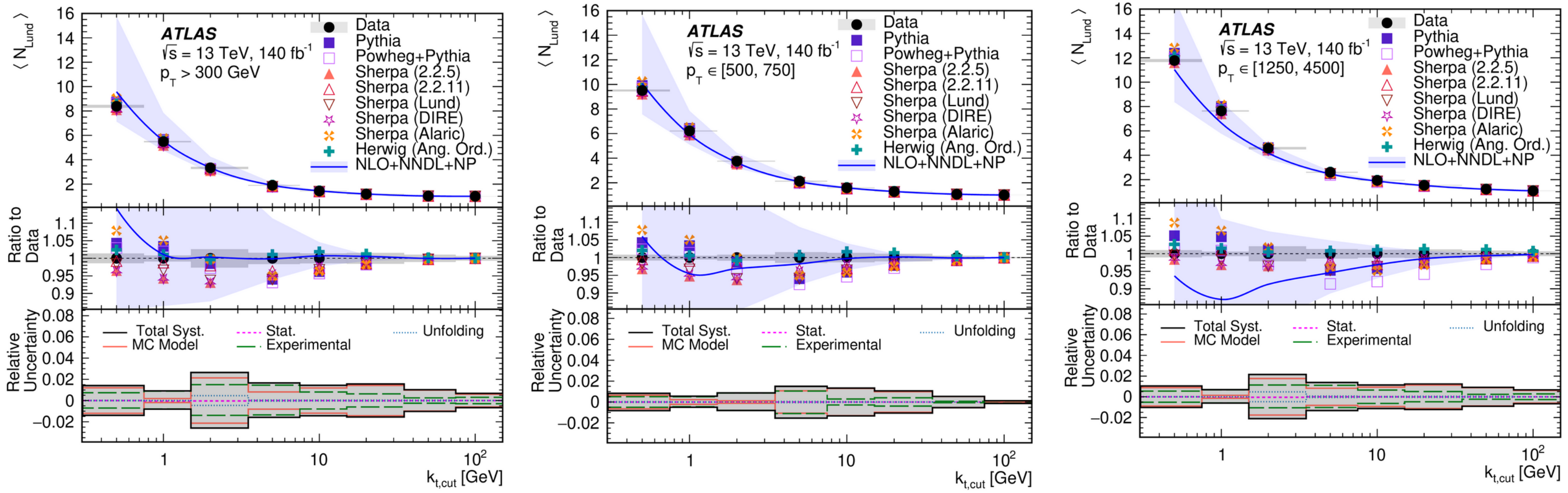
Average Lund multiplicity

Average multiplicity <N_{Lund} > against k_t cut-off in pT bins

➔ **HERWIG Angular Ordered best**

Analytic prediction NLO+NNDL reweighted to include non-perturbative effects (NP)

- Agrees well with MC and data
- At low k_t cut-off , **increasing disagreement with p_T** : worse than MC



Jet pT increases

Improving Monte Carlo predictions

ATLAS result in dijet events from Run 2
arXiv:2402.13052
*Measurements of **Lund subjet multiplicities** in 13 TeV proton-proton collisions with the ATLAS detector*

ATLAS result in ttbar events from Run2
Phys. Rev. D 109 (2024) 112016
*Measurement of **jet substructure in boosted $t\bar{t}$ events** with the ATLAS detector using 140 fb^{-1} of 13 TeV pp collisions*

ATLAS result in Z+jets events from Run 2
arXiv:2405.20041
*A **simultaneous unbinned differential cross section** measurement of twenty-four Z+jets kinematic observables with the ATLAS detector*

Unbinned simultaneous unfolding

Flaws of Iterative Bayesian Unfolding:

- Define target observables
- Define binning
- 1D, eventually 2D or 3D (but statistically limited)

➔ Mismodelling due to hidden variable

Goal here: unfold **simultaneously** 24 observables at event level in an **unbinned** way

➔ OMNIFOLD method

In Z+jets events
With large hadronic recoil

Event selection	
Trigger	Single muon trigger
Muon selection	2 opposite charge, $p_{\mathrm{T}} > 25\,\mathrm{GeV}$, $\lvert\eta\rvert < 2.5$
Z candidate	$m_{\mu\mu} \in [81; 101]\,\mathrm{GeV}$
Transverse hadronic recoil	$p_{\mathrm{T}}^{\mu\mu} > 200\,\mathrm{GeV}$
Jet selection	Only the 2 leading p_{T} jets
Jet definition	
Constituents	Charged tracks (except the 2 muons)
Clustering	Anti-k_{T} $R = 0.4$
Backgrounds	
Top processes ($t\bar{t}, tW$)	Treated as an uncertainty

Event $\vec{\mathbf{x}} =$

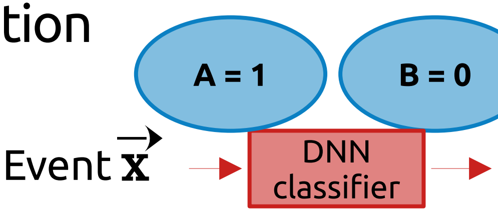
List of variables	
Z production	$p_{\mathrm{T}}^{\mu\mu}$, $y_{\mu\mu}$
μ kinematics	both muons p_{T}^{μ}, η_{μ}, ϕ_{μ}
Jets kinematics	both jets p_{T}^{j}, η_{j}, ϕ_{j}
Jets substructure	both jets masse m_j, charged multiplicity n^{ch} N-subjettiness τ_1, τ_2, τ_3

Large overlap with Pythia A14 tune variables
→ ***Improve tuning !***

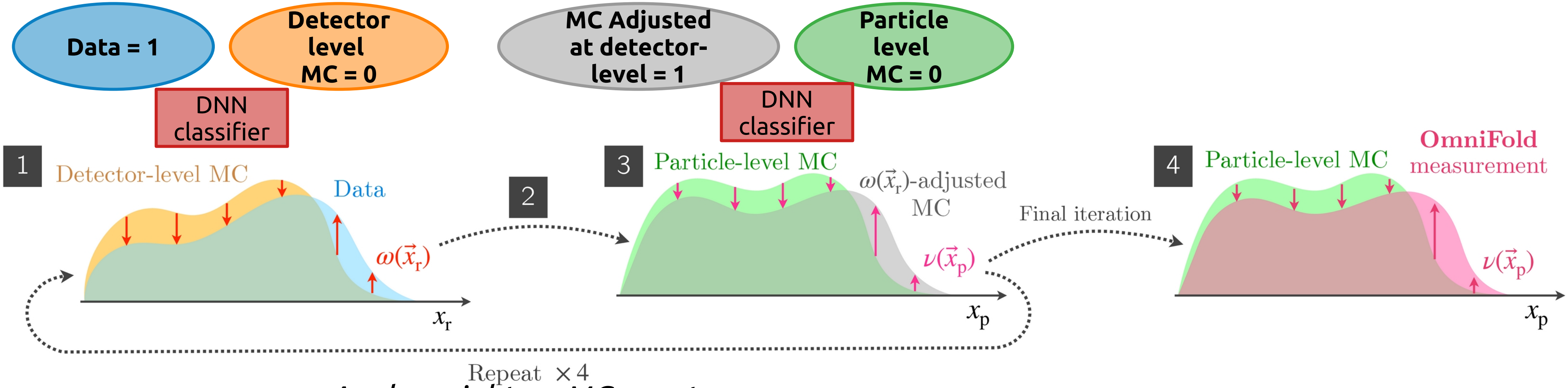
The OMNIFOLD method

DNN classifier score used to reweight a distribution

- **Reweight successively** the distribution at detector level and particle level
- Each reweigting needs **a new trained DNN**



$$\frac{\mathrm{DNN}(\vec{x})}{1-\mathrm{DNN}(\vec{x})} \approx \frac{\mathbb{P}(A)}{\mathbb{P}(B)} = weight(\vec{x})$$



Apply weight on MC event
→ both detector and particle level affected

Unfolded dataset published

Unbinned unfolding:

– Free choice of binning*

– Combination of variables possible

– Sub-region selection

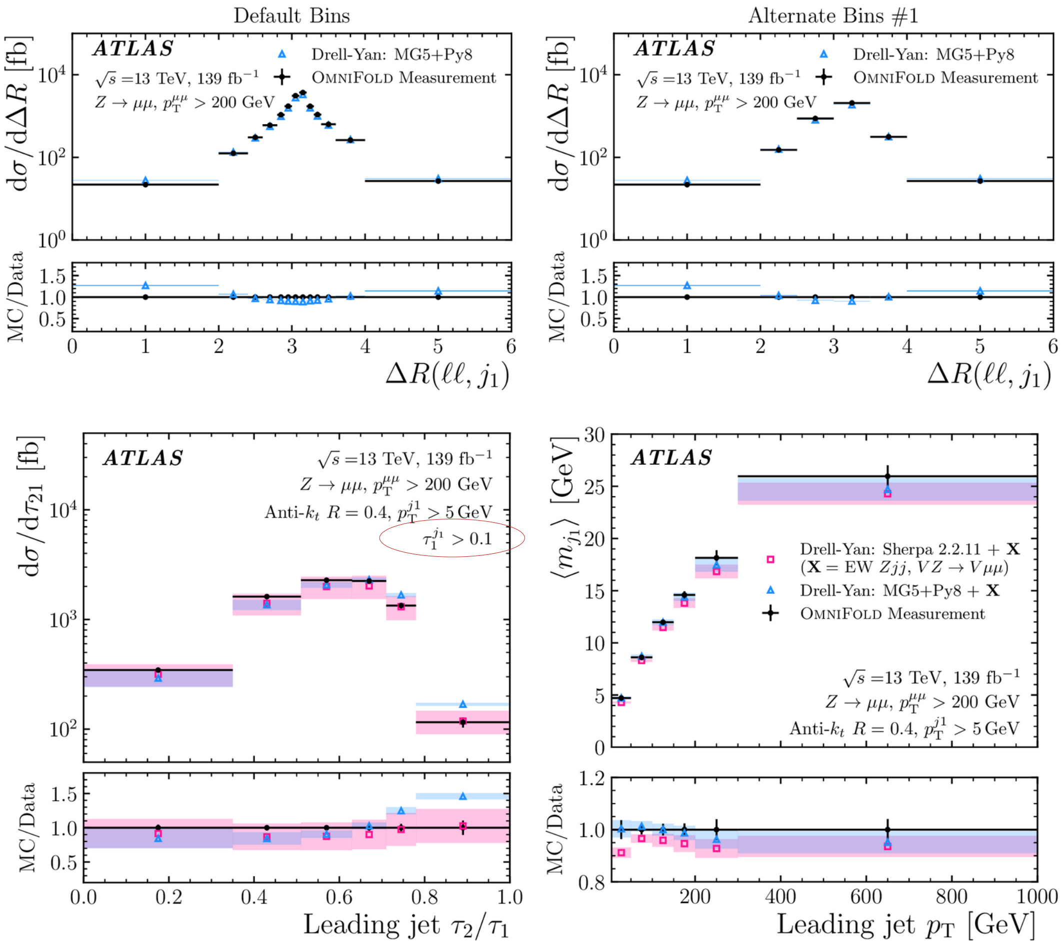
➔ **Flexibility !**

First unfolded unbinned dataset published at the LHC

– OMNIFOLD = rather lightweight and ready-to-use unfolding tool

– Asymptotically reduces to IBU

➔ Template for **future such unfolding !**



** Up to certain best practices to maintain statistical power in the bins*

Study of jet fragmentation

ATLAS result in dijet events from Run2 :

ATLAS-CONF-2024-012
*Measurement of **Track Functions** in ATLAS Run 2 Data*

CMS results in jets from low pile-up pp run before heavy ions (301 pb^{-1})

CMS-PAS-HIN-24-005
*Jet fragmentation function and groomed substructure of **bottom quark jets** in proton-proton collisions at 5.02 TeV*

CMS-PAS-HIN-24-007
*Exploring small-angle emissions in prompt **D^0 jets** in proton-proton collisions at $\sqrt{s}$= 5.02 TeV*

Observable definition

Goal: Study **non-perturbative QCD** in jets

– Fragmentation function for partons in **single** hadron

– Lack correlations between hadrons in reverse process: **hadronisation**

➔ Jet level observable: track function

$$r_q = \frac{p_T^{\text{charged}}}{p_T^{\text{all}}}$$

Study the **moments** of its distribution

– First moment **~ 2/3** (QCD approximate Isospin symmetry)

– Higher moments encode **correlations in hadronisation**

In dijet events

Jet definition	
Constituents	Particle Flow objects
Clustering	Anti-k_T $R = 0.4$
Event selection	
Trigger	Single jet trigger
Jet kinematic	$p_T^{jet} > 240$ GeV and $\|y\| < 2.1$
Dijet balance	$p_T^{leading} < 1.5 \times p_T^{subleading}$
Jet selection	Only the 2 leading jets
Backgrounds	
Non purely $2 \to 2$ QCD processes : *top quark, vector bosons,...*	Negligible

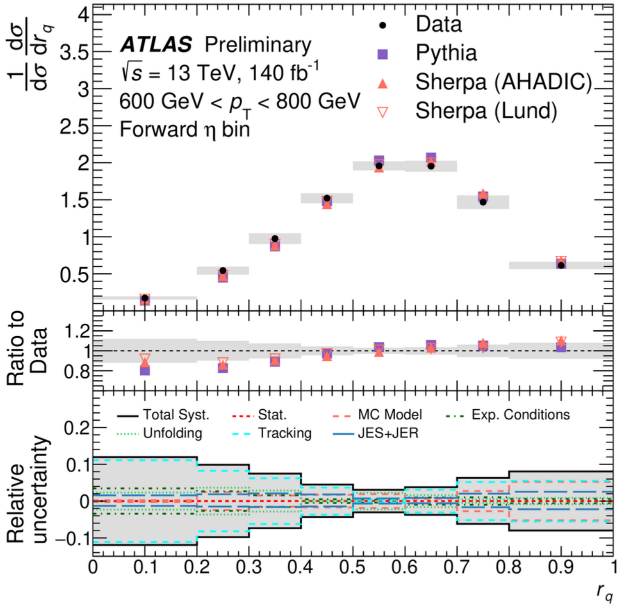
Unfolded distribution

Iterative Bayesian Unfolding of **r_q, p_T** and **$\eta_{\text{central vs forward}}$**

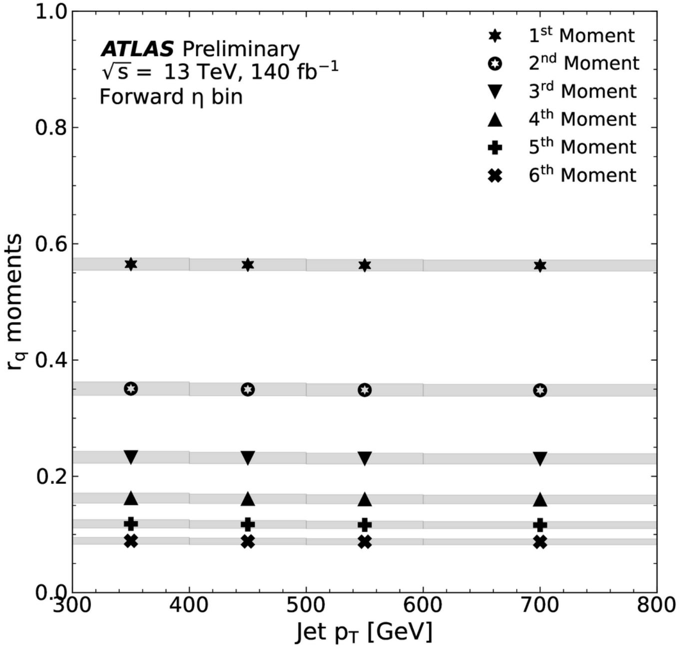
➔ Uncertainties at 5 to 12 %, dominated by tracking uncertainties

Up to 6th moment extracted in bins of pT

➔ Very slowly evolving with jet p_T



x pT bins



From moments

$$\mu_n = \mathrm{E}[(X - \mathrm{E}[X])^n]$$

to cumulants

$$K(t) = \log \mathrm{E}[e^{tX}] = \sum_{n=1}^{\infty} \kappa_n \frac{t^n}{n!}$$

$$\kappa_2 = \mu_2\,,$$

$$\kappa_3 = \mu_3\,,$$

$$\kappa_4 = \mu_4 - 3\mu_2^2\,,$$

$$\kappa_5 = \mu_5 - 10\mu_3\mu_2\,,$$

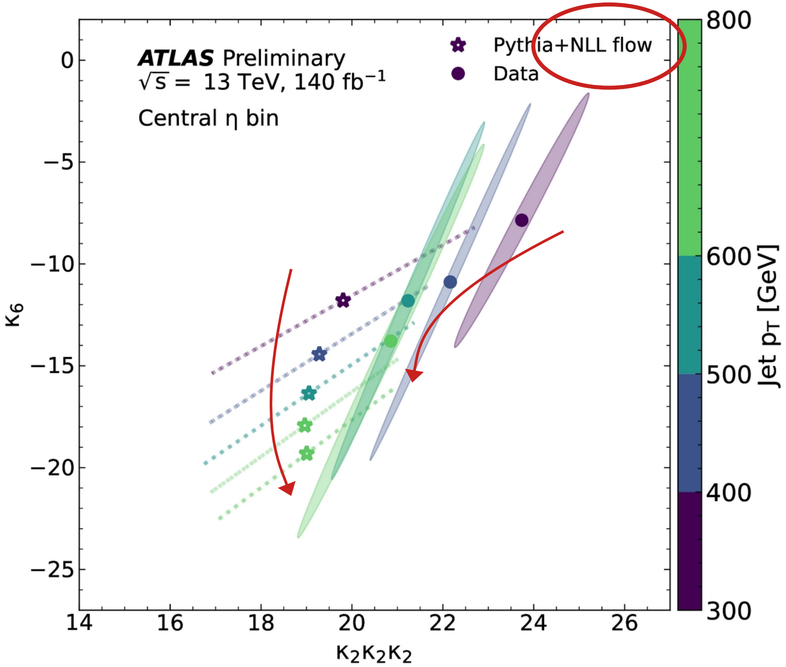
$$\kappa_6 = \mu_6 - 15\mu_4\mu_2 - 10\mu_3^2 + 30\mu_2^3\,.$$

The cumulant representation

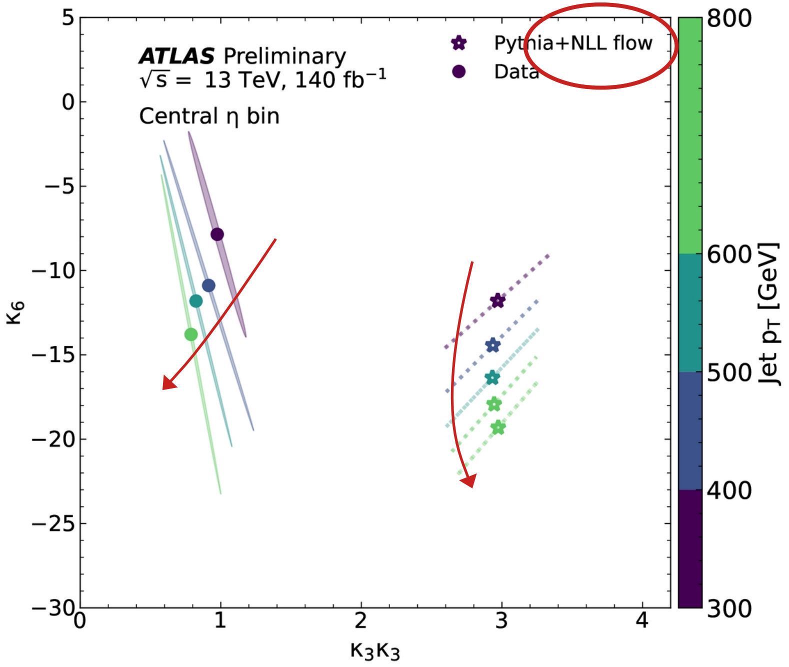
Track function have a **non-linear renormalisation group flow**
as opposed to DGLAP regime of standard fragmentation function

Flow in the cumulant representation

➔ **Should converge to a fixed point**



Data and MC converging



Data and MC diverging !

Study of jet fragmentation

ATLAS result in dijet events from Run2 :

ATLAS-CONF-2024-012

*Measurement of **Track Functions** in ATLAS Run 2 Data*

CMS results in jets from low pile-up pp run before heavy ions (301 pb^{-1})

CMS-PAS-HIN-24-005

*Jet fragmentation function and groomed substructure of **bottom quark jets** in proton-proton collisions at 5.02 TeV*

CMS-PAS-HIN-24-007

*Exploring small-angle emissions in prompt **D^0 jets** in proton-proton collisions at $\sqrt{s}$= 5.02 TeV*

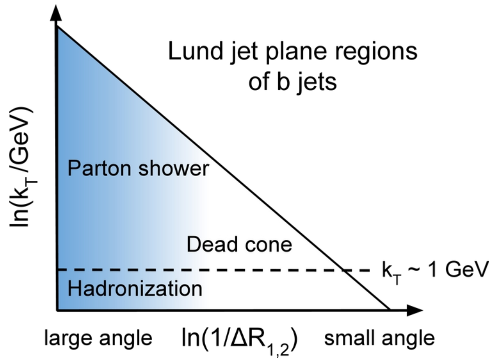
Motivation

Heavy flavour jet fragmentation from DGLAP evolution:

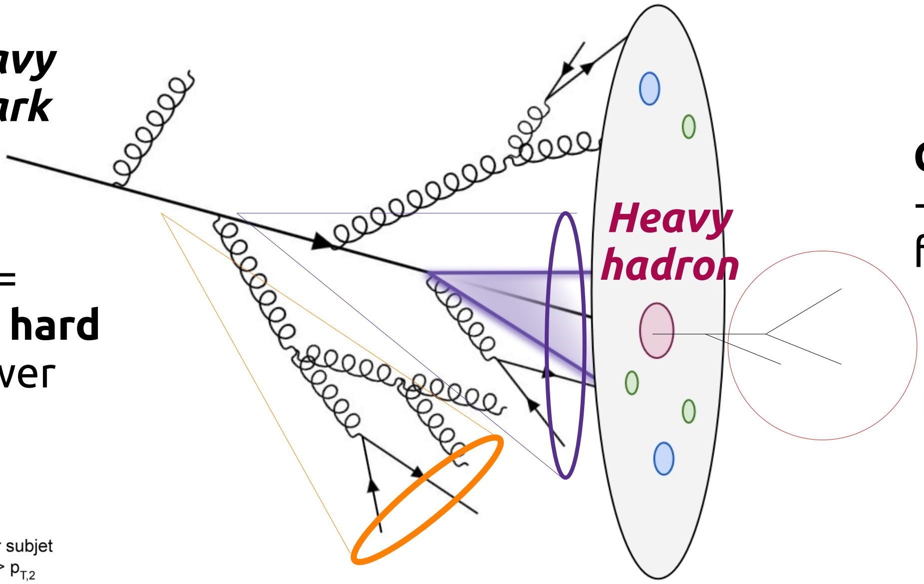
$$P_{(Q\to Qg)}(z) = \frac{1-z}{z} + \frac{z}{2} - 2\frac{m_{\mathrm{Q}}^2}{m_{\mathrm{Qg}}^2 + m_{\mathrm{Q}}^2}$$

Dead-cone effect →

$$d\mathcal{P}(\theta) \propto \frac{d\theta^2}{(\theta^2+\theta_0^2)^2} \text{ with } \theta_0 = m_{\mathrm{Q}}/E_{\mathrm{Q}}$$



Heavy quark



Jet grooming
→ Split in **2 subjets** = investigate the **first hard splitting** in the shower

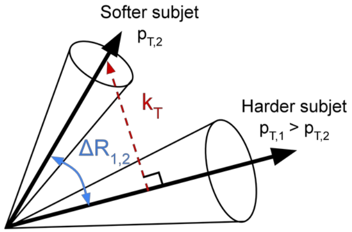
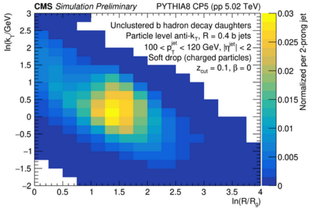


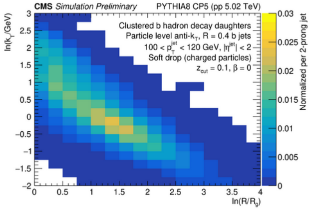
Figure by Lida Kalipoliti

Group hadron decay
→ separate hadron decay from Parton shower

Lund Jet Plane simulations



Unclustered b-hadron



Clustered b-hadron

Analysis method

Select an **inclusive** population of jets
and

b-jets

Event selection

Trigger	single jet triggers at low p_T
Inclusive Jet selection	$100 < p_T < 120\,\mathrm{GeV}$, $\lvert\eta\rvert < 2.0$
b-jet	GNN tagger at 52 % efficiency 99.9 % purity

Inclusive jets

b-jets

D^0 jets

Event selection

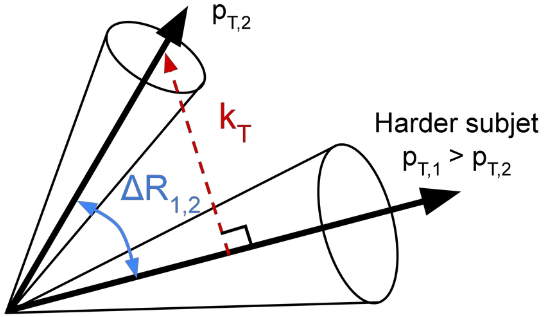
Trigger	single jet triggers at low p_T
Inclusive Jet selection	$100 < p_T < 120\,\mathrm{GeV}$, $\lvert\eta\rvert < 1.6$
D^0 meson candidate	2 oppositely charged tracks with $m^{inv} = 1.86 \pm 0.2\,\mathrm{GeV}$, $p_T^{inv} > 4GeV$, $\lvert y^{inv}\rvert < 1.2$
D^0 jet	$\Delta R(jet, D^0) < 0.2$
non-prompt rejection	Distance of Closest Approach $DCA/DCA_{err} < 4$

D^0 jets

Hadron decay: Machine learning classifier to cluster b-hadron decay tracks

Jet grooming
→ Soft-Drop

Unfold **ΔR_g**, p_T fraction **z_g** and b-quark p_T fraction **z_b**



More details in back-up

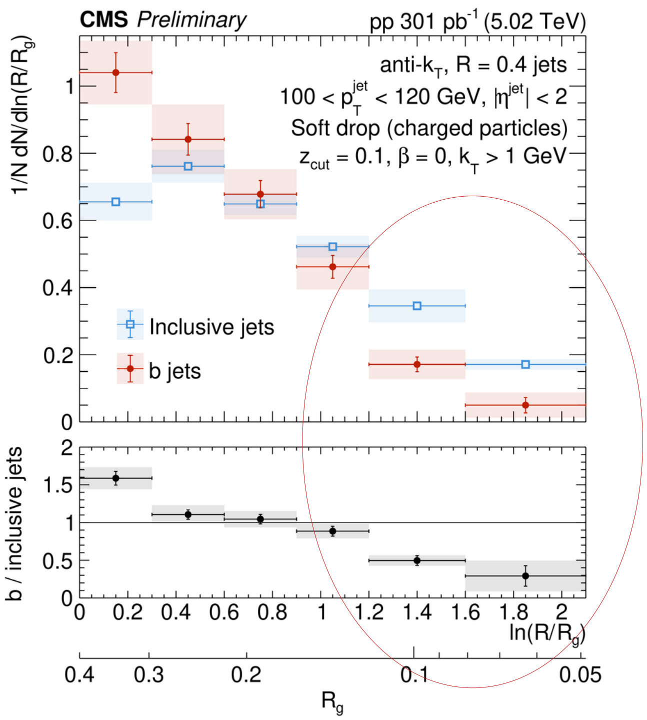
Hadron decay: remove $D^0 \rightarrow K\,\pi$ decay tracks

Jet grooming
→ Late-k_T

Unfold angle **$\theta_g \sim \Delta R_g$**

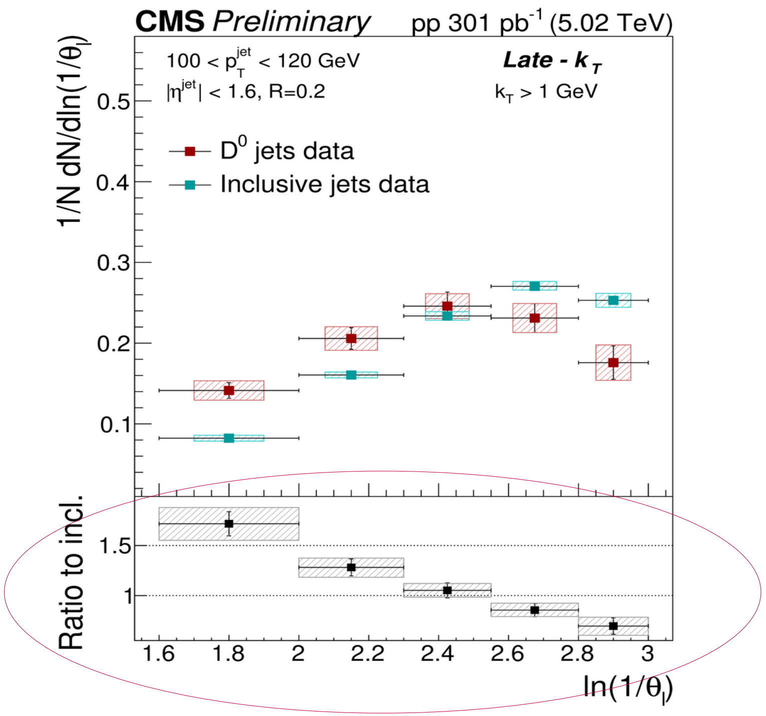
$$k_T = \theta_g\, p_T^{emission}$$

Dead-cone effect



In b-jets

➔ Supression at low R = large ln(R/Rg)



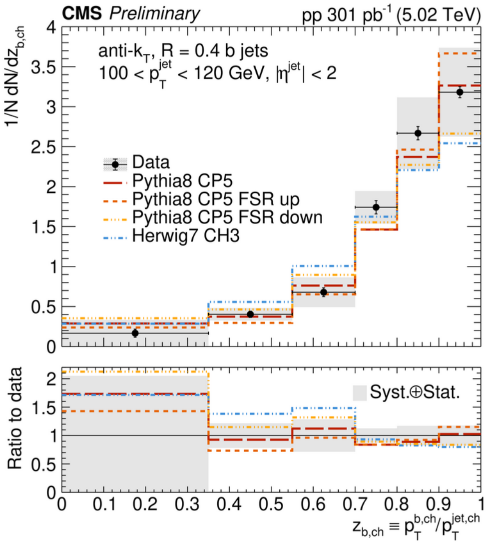
In D^0-jets

➔ Supression at low θ = large ln(1/θ)

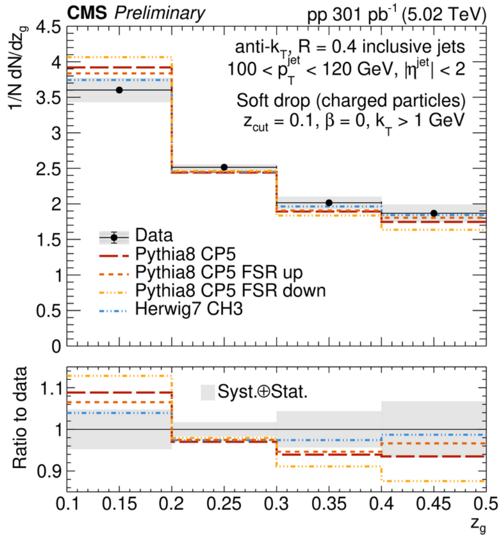
➔ Most visible with late-k_T grooming

Modelling of heavy quark jets

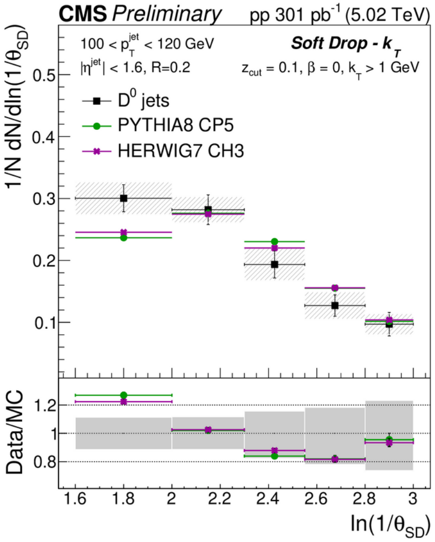
b jets



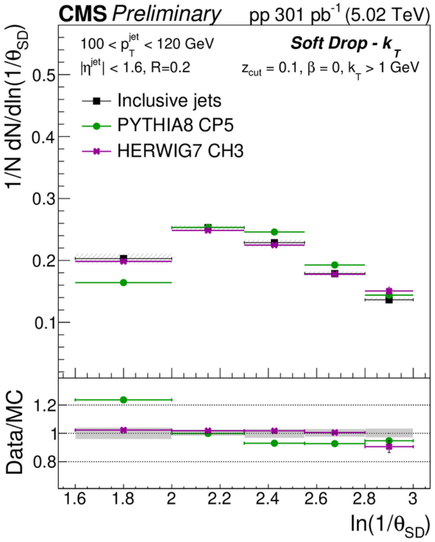
Inclusive jets



D⁰ jets



Inclusive jets



In b-jets

- **Correct** b-jet modelling by **Pythia**
- **Mismodelling in Herwig** itself better in inclusive

In D^0 jets

- **Herwig** better in inclusive jets
- Similar **Mismodelling in Herwig** and **Pythia** for D^0 jets

The Lund Jet Plane and other substructure variables studies can

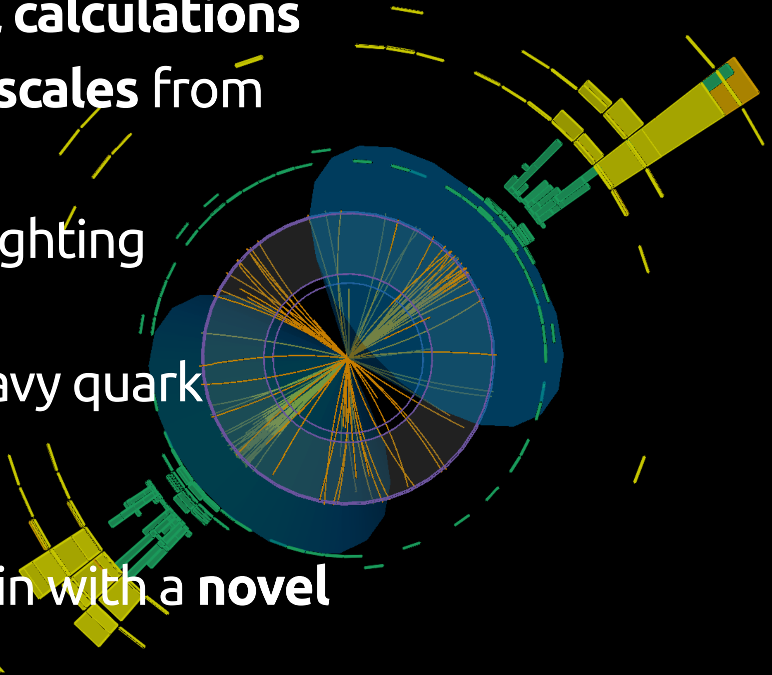
- Test and improve **state-of-the-art theoretical calculations**

➔ Modelling of QCD in a **wide range of energy scales** from hadronisation to perturbative regime

- Provide **data-driven** tools for tagging or reweighting

CMS provides insight in the **dead-cone effect** in heavy quark fragmentation

ATLAS publishes the first particle-level dataset obtain with a **novel unbinned unfolding** method : **OMNIFOLD**



Thank you for your attention

Extra

QCD jets Lund plane: events selection & unfolding

JHEP 05 (2024) 116 (CMS)

Jet definition	
Constituents	Particle Flow objects
Clustering	Anti-k_T $R = 0.4$ or 0.8
Event selection	
Trigger	Single jet trigger
Jet $\|y\|$	< 1.7
Jet p_T [GeV]	$p_T^{jet} > 700$
Backgrounds	
Non purely QCD processes : *top quark, vector bosons,...*	Highly negligible

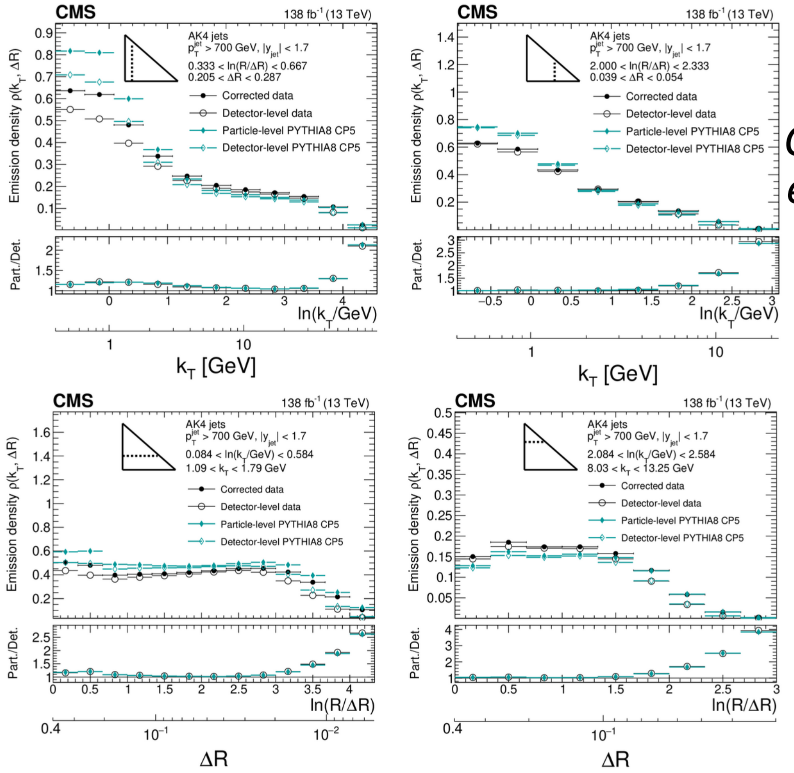
Lund jet plane **unfolded** (Iterative Bayesian Unfolding) separately for R=0.4 and R=0.8 jets

Uncertainties dominated by parton shower and hadronisation modelling (2-7%)

From *detector level* to **particle level**

Parton cascade region

Collinear emission region



Hadronisation and UE region

Hard emission region

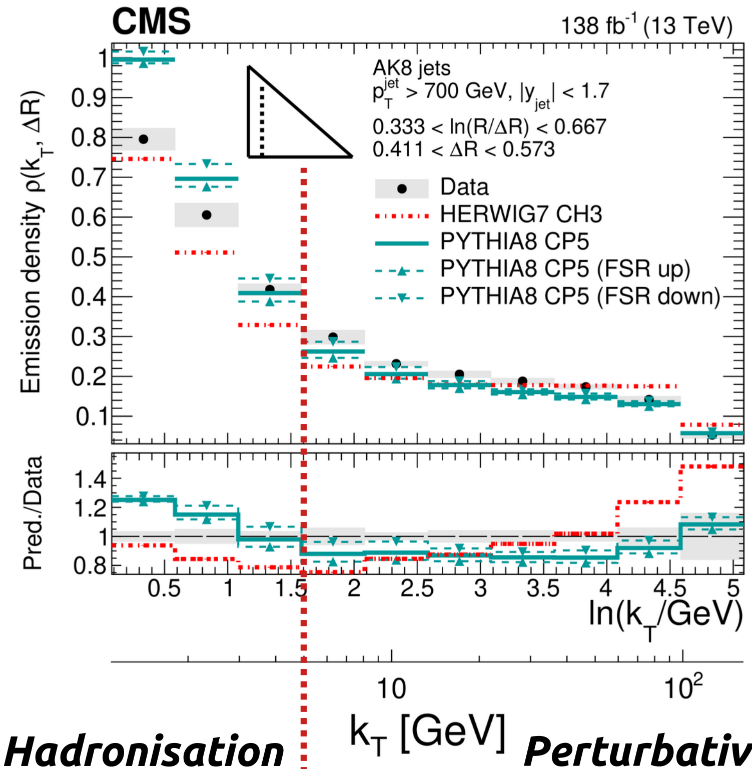
LJP Factorises effects : FSR

JHEP 05 (2024) 116 (CMS)

α_s^{FSR} variations decreases with decreasing k_T

– FSR Down = α_s^{FSR} increased favoured by data

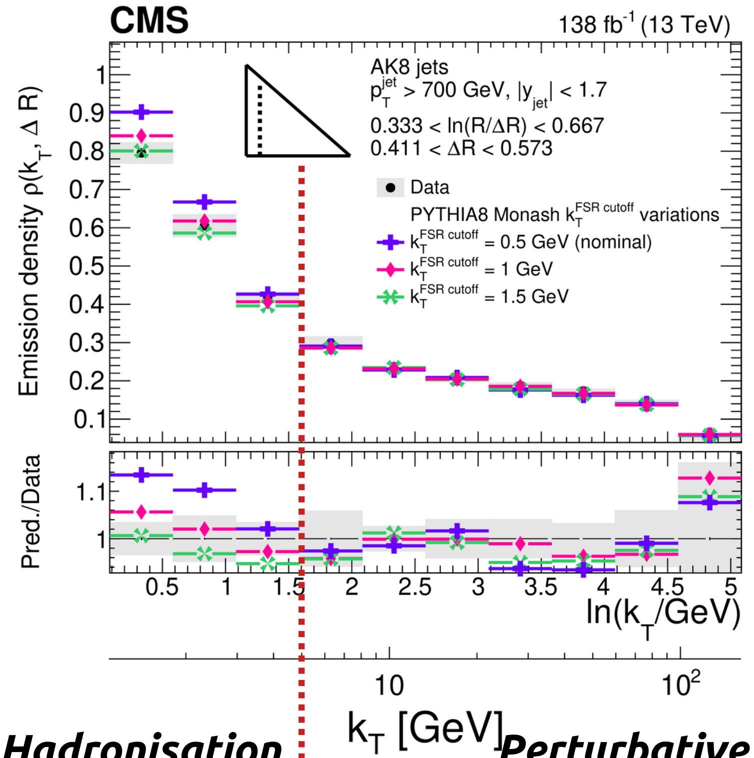
➔consistent with ttbar !



Hadronisation effect

Perturbative radiations

Impact of cut-off scale to start hadronisation



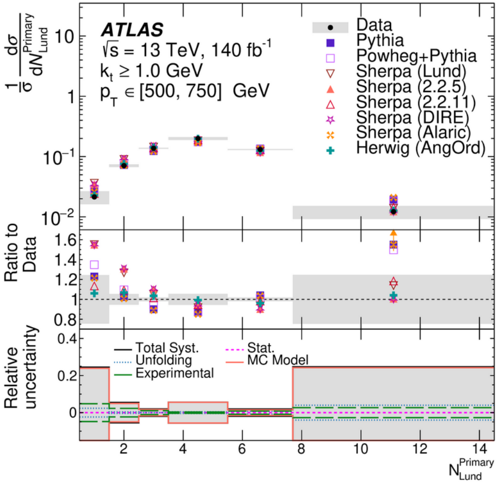
Hadronisation effect

Perturbative radiations

Event selection & unfolding of Lund subjet multiplicity

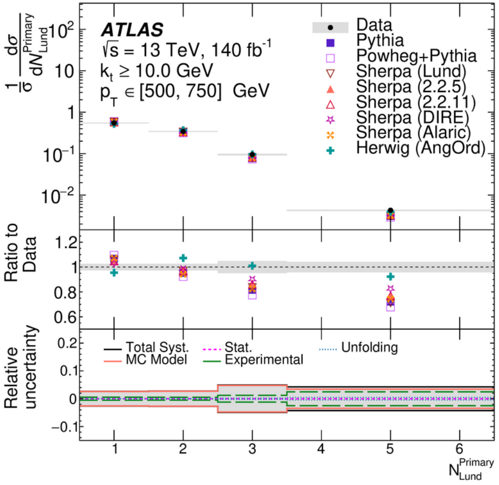
arXiv:2402.13052 (ATLAS)

Jet definition	
Constituents	Particle Flow objects
Clustering	Anti-k_T $R = 0.4$
Event selection	
Trigger	Single jet trigger
Jet kinematic	$p_T^{jet} > 120$ GeV and $\lvert y \rvert < 2.1$
Dijet balance	$p_T^{leading} < 1.5 \times p_T^{subleading}$
Jet selection	Only the 2 leading jets
Backgrounds	
Non purely $2 \to 2$ QCD processes : *top quark, vector bosons,...*	Negligible



N_{Lund} and $N_{Lund}^{Primary}$ at particle level with an **Iterative Bayesian Unfolding**

➔ In bins of p_T, $\eta_{\text{central vs forward}}$, and with varied k_t cut-off



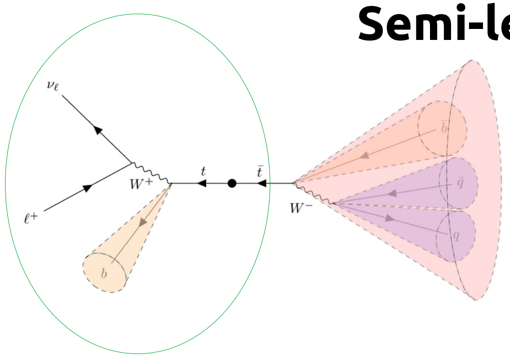
Main uncertainties at 2% – 10 %

- Dominated by MC modelling
- Shrink with increasing cut-off

Studying Large-R jets from top quark decay

	Jet definition	
	All-hadronic channel	***Semi-leptonic channel***
Constituents	Weighted calorimeter topo clusters	$R = 0.4$ Particle Flow jets
Clustering	Anti-k_T $R = 1.0$	
Grooming	Trimming $R_{trim} = 0.2$, $f_{trim} = 0.05$	

Semi-leptonic channel

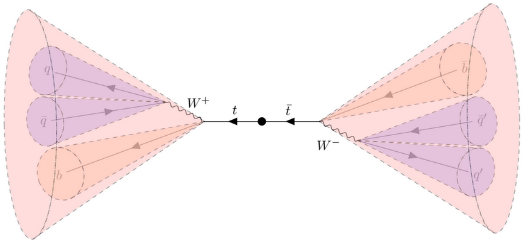


Event selection	
Trigger	Single lepton ($e^{\pm}, \mu^{\pm}$) trigger
W Lepton ℓ	$= 1$ passing lepton selection
W Neutrino	$E_T^{miss} > 20$ GeV, $E_T^{miss} + m_T^W > 60$ GeV
Small-R jets	≥ 3, $p_T > 30$ GeV, $\lvert\eta\rvert < 2.5$
b-tagged jet j_b	≥ 1 and $\Delta R(\ell, j_{b1}) < 1.5$
Large-R Jet J	$p_T^{jet} > 350$ GeV, $\lvert y\rvert < 2.0$, $N_{\text{Small-R jet}} \geq 2$
Top candidate	$m_J \in [m_{top} - 50\,\text{GeV}; m_{top} + 50\,\text{GeV}]$, $\Delta R(\ell, J) > 1.0$
b-jet closest to ℓ	not a constituent J, $m(\ell, j_b) < 120$ GeV
Backgrounds	
Irreducible tW, V+jets $t\bar{t}V$, $t\bar{t}H$, VV	MC simulations
Fake leptons	Data-driven *Matrix Method*

Goal: study in **boosted** regime the substructure of top jets

➔ Average jet **pT ~ 500GeV**

All-hadronic channel



Event selection	
Trigger	Collection of single and double Large-R jets triggers
Lepton veto	$= 0$ lepton with $p_T > 25$ GeV
Large-R Jet J	≥ 2 with $p_T^{jet} > 350$ GeV, ≥ 1 with $p_T^{jet} > 500$ GeV, $\lvert\eta\rvert < 2.0$
Top mass	$m_J \in [m_{top} - 50\,\text{GeV}; m_{top} + 50\,\text{GeV}]$
b-matching	2 leading J ghost-associated to a b-tagged track jet at 70%
Top-tagging	≥ 1 of the 2 leading J top tagged at 80%
Backgrounds	
$t\bar{t}$ semi-leptonic contamination	MC simulations
Irreducible tW, V+jets $t\bar{t}V$, $t\bar{t}H$, VV	MC simulations
Multijet production	Data-driven : generalised ABCD method

Orthogonality

Only jets with pT > 500 GeV are studied

Two taggers for both jets
→ 16 regions

2nd large-R jet \ 1st large-R jet	t0b0	t1b0	t0b1	t1b1
t1b1	J	K	L	S
t0b1	B	D	H	N
t1b0	E	F	G	M
t0b0	A	C	I	O

Substructure variables unfolding

Phys. Rev. D 109 (2024) 112016 (ATLAS)

8 substructure variables from “ghost associated” tracks unfolded to particle level using **Iterative Bayesian Unfolding** in both channels

Angularities	$p_{\mathrm{T}}^{\mathrm{d},*} = \sqrt{\left(\sum_{i\in J} z_i^2 - \frac{1}{N}\right)\frac{N}{N-1}}$	Momentum distribution
	$\mathrm{LHA} = \sum_{i\in J} z_i \sqrt{\frac{\Delta R(i,J)}{R}}$	Broadness of the jet
Energy correlation functions	$D_2 = \frac{ECF(3)ECF(1)^3}{ECF(2)^3}$	Two-prong structure
	$C_3 = \frac{ECF(4)ECF(2)}{ECF(3)^2}$	Three prong structure
	$\mathrm{ECF2} = \frac{\mathrm{ECF}(2)}{\mathrm{ECF}(1)^2}$	Linked to both C_3 and D_2
N-subjettiness	$\tau_{21} = \frac{\tau_2}{\tau_1}$	Two-prong structure
	$\tau_{32} = \frac{\tau_3}{\tau_2}$	Three-prong structure
	τ_3	Poorly modelled in MC

$$\mathrm{ECF}(1,\beta) = \sum_{i\in J} p_{T\,i},$$
$$\mathrm{ECF}(2,\beta) = \sum_{i<j\in J} p_{T\,i}\, p_{T\,j} (R_{ij})^{\beta},$$
$$\mathrm{ECF}(3,\beta) = \sum_{i<j<k\in J} p_{T\,i}\, p_{T\,j}\, p_{T\,k} (R_{ij}R_{ik}R_{jk})^{\beta},$$
$$\mathrm{ECF}(4,\beta) = \sum_{i<j<k<\ell\in J} p_{T\,i}\, p_{T\,j}\, p_{T\,k}\, p_{T\,\ell} (R_{ij}R_{ik}R_{i\ell}R_{jk}R_{j\ell}R_{k\ell})^{\beta}.$$

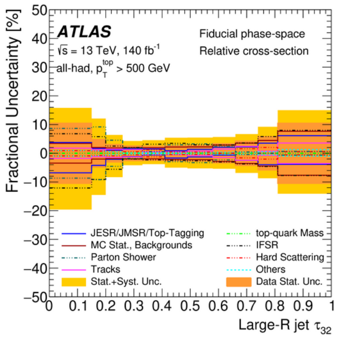
$$\tau_N = \frac{1}{d_0}\sum_k p_{\mathrm{T},k} \min\{\Delta R_{1,k}, \Delta R_{2,k}, \ldots, \Delta R_{N,k}\}$$
$$\text{with} \quad d_0 = \sum_k p_{\mathrm{T},k} R_0.$$

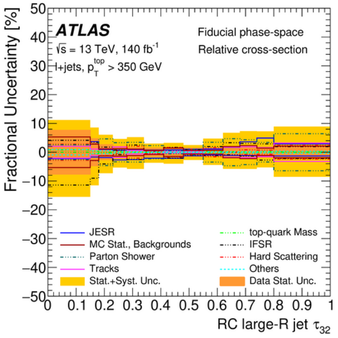
Degree to which a jet is compatible with having ≤ N subjets

*Most **sensitive** variables in **W-jet/top-jet multivariate discriminator** used by ATLAS*

Uncertainties at 5% in most bins, increases at tails of distributions

➔Main uncertainties are theoretical: **FSR**, ***parton shower and hadronisation*** modelling





Analysis method

B-jet fragmentation — CMS-PAS-HIN-24-005

Jet definition	
Constituents	Particle Flow objects
Clustering	Anti-k_T $R = 0.4$
Grooming	Soft-drop $z_{cut} = 0.1, \beta = 0, k_T > 1\,\mathrm{GeV}$
Event selection	
Trigger	single jet triggers at low p_T
Inclusive Jet selection	$100 < p_T < 120\,\mathrm{GeV}$, $\lvert\eta\rvert < 2.0$
b-jet	GNN tagger at 52 % efficiency 99.9 % purity
Backgrounds	
light+c jets	MC simulation
double-b jets ($g \to bb$)	MC template fit of $m_{b\ had.}$

B-hadron decay **clustered by a machine learning** classifier

Unfolded distributions

– **Soft-drop grooming**: $\mathbf{R_g}$ and $\mathbf{z_g}$

– B-hadron p_T fraction $z_{b,ch} = \frac{p_T^{b,ch}}{p_T^{jet,ch}}$

D^0-jet fragmentation — CMS-PAS-HIN-24-007

Jet definition	
Constituents	Particle Flow objects
Clustering	Anti-k_T $R = 0.2$
Grooming	Late-k_T with $k_T > 1\,\mathrm{GeV}$ Soft-drop $z_{cut} = 0.1, \beta = 0, k_T > 1\,\mathrm{GeV}$
Event selection	
Trigger	single jet triggers at low p_T
Inclusive Jet selection	$100 < p_T < 120\,\mathrm{GeV}$, $\lvert\eta\rvert < 1.6$
D^0 meson candidate	2 oppositely charged tracks with $m^{inv} = 1.86 \pm 0.2\,\mathrm{GeV}$, $p_T^{inv} > 4GeV$, $\lvert y^{inv}\rvert < 1.2$
D^0 jet	$\Delta R(jet, D^0) < 0.2$
non-prompt rejection	Distance of Closest Approach $DCA/DCA_{err} < 4$
Backgrounds	
Combinatorial background	Power law in m_{D^0} template fit
non-promtp D^0 subtraction	MC template fit of DCA/DCA_{err}

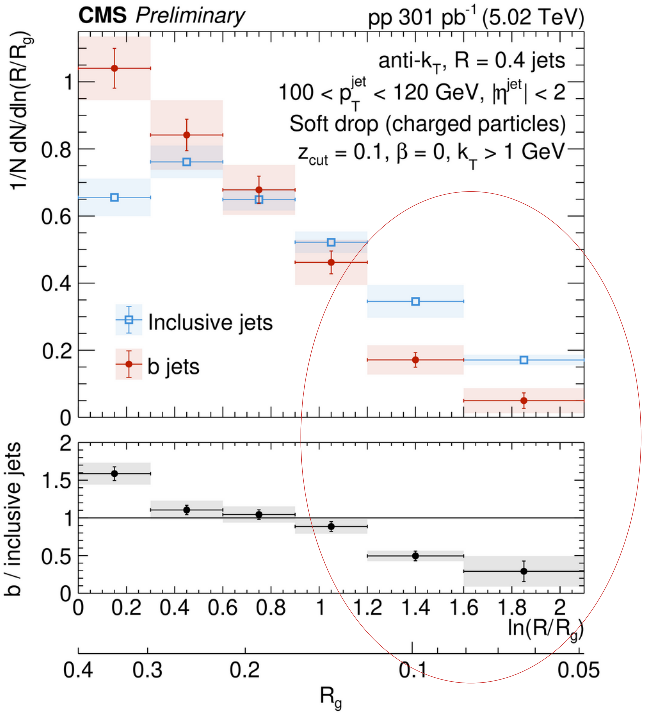
D^0 Meson **decay tracks removed** from the grooming algorithm

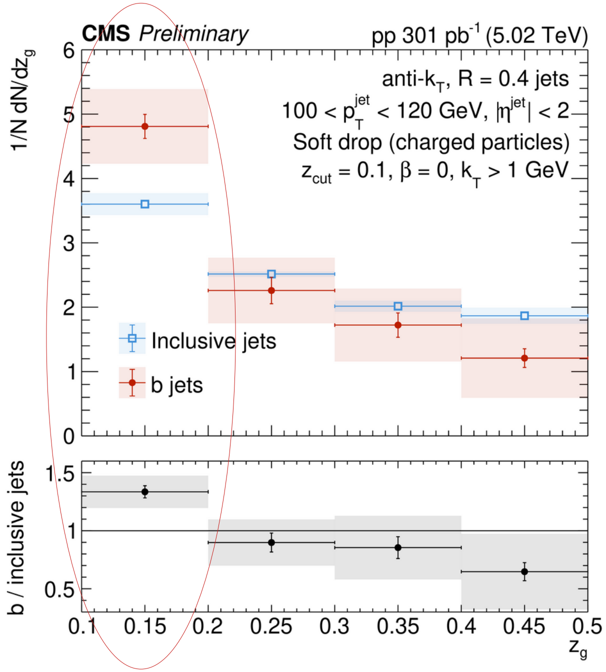
2 Grooming angles unfolded $k_T = \theta_g\, p_T^{emission}$

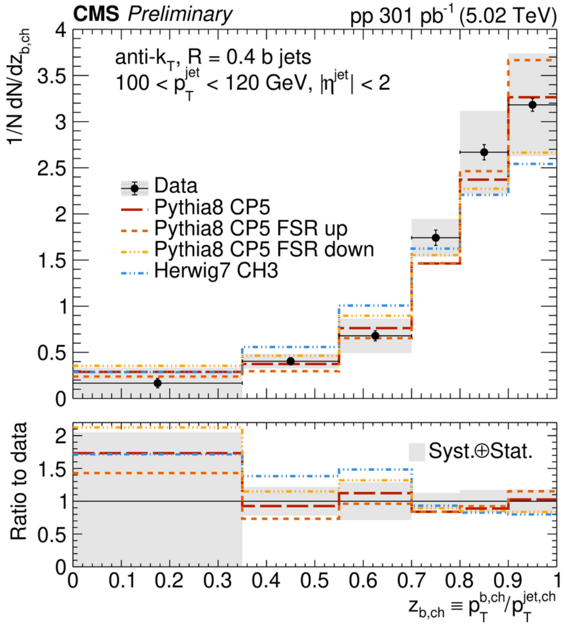
– **Late-$\mathbf{k_T}$**: sensitive to hard collinear emission ~ c → cg splitting

– **Soft-Drop**: sensitive to larger angle emission ~ g→ cc splitting

Results in b-jets







Dead-cone effect in b-jets

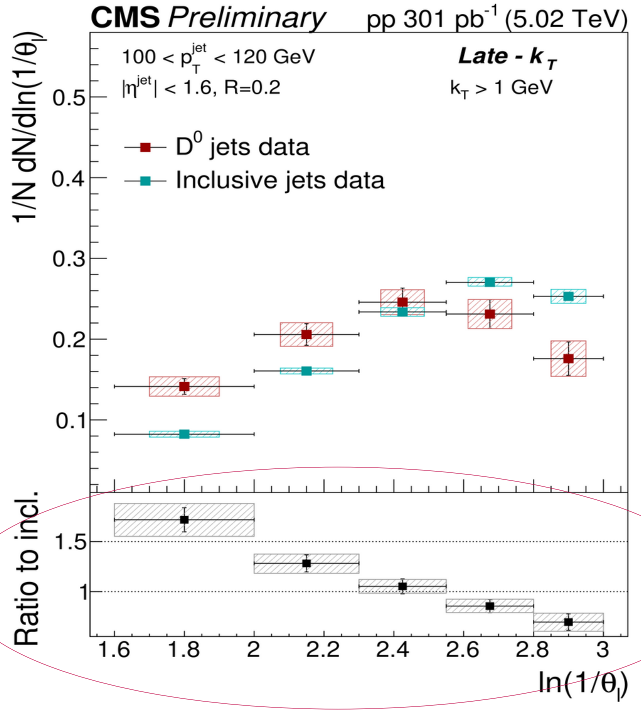
➔ Supression at low R
= large ln(R/Rg)

Higher fraction of splitting momentum taken by b-quark

– **Correct** b-jet modelling by **Pythia**

– **Mismodelling in Herwig** (itself better in inclusive)

Results in D^0 jets

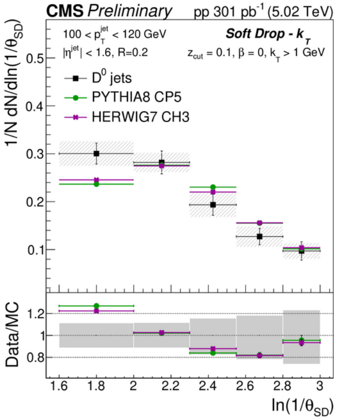


Dead-cone effect in D^0-jets

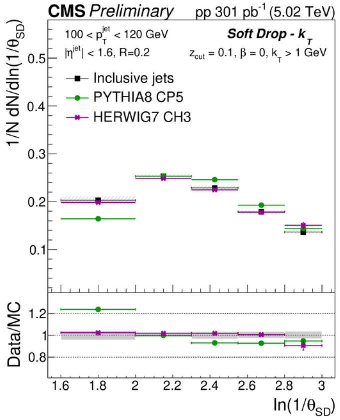
➔ Supression at low θ = large ln(1/θ)

➔ Most visible with late-k_T grooming

D^0 jet



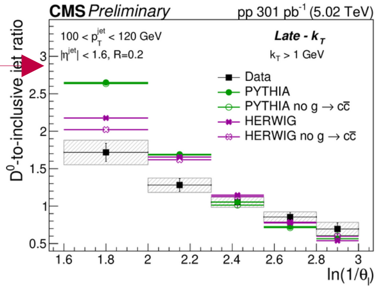
Inclusive jet



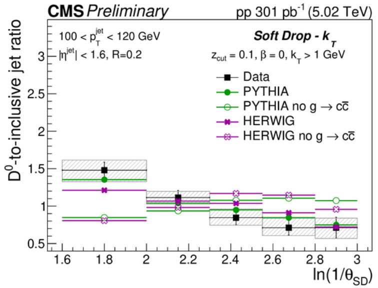
– **Herwig** better in inclusive jets

– Similar **Mismodelling in Herwig** and **Pythia** for D^0 jets

Late-k_T



Soft Drop



Late-kT modelling resilient to **g→ cc**

➔Soft-Drop mismodelled *without g→cc*