Flavoured jets and how to define them

Rene Poncelet

THE HENRYK NIEWODNICZAŃSKI
INSTITUTE OF NUCLEAR PHYSICS
POLISH ACADEMY OF SCIENCES

Flavoured jets are everywhere

- Flavoured jets as signature
  - Top-quarks
  - Vector+heavy flavour: pp → W/Z/A + c/b
  - Higgs → charm, Higgs → bottom
  - New physics searches
  - Studies of QCD dynamics

Rely on our capability to
→ identify (i.e. tag) flavoured jets
→ interpret (i.e. predict) them

- Focus of this talk: V + heavy-flavour (→ but many aspects are generalisable)
  - Benchmark for flavour tagging
  - IR safety/sensitivity

Outline

- What do we mean by "flavoured jets" and why are there problems?
- Anti-kT “like” flavoured jet algorithms
- Phenomenology → Les Houches study
- Interface to experiment
- Wrap up & outlook

A look back: Snowmass accord 1990 …

Toward a Standardization of Jet Definitions

John E. Huth and Naor Wainer
Fermi National Accelerator Laboratory
P.O. Box 500
Batavia, Illinois 60510

Karlheinz Meier
Deutsches Elektronen Synchrotron (DESY)
Hamburg 52, Germany

Nicholas Hadley
University of Maryland
College Park, Maryland 20742

F. Aversa and Mario Greco
Instituto Nazionale di Fiscia Nucleare (INFN)
Frascati, Italy

P. Chiappetta and J. Ph. Guillet
CTP-CRNS, Luminy
Marseille, France

Stephen Ellis
University of Washington
Seattle, Washington 98195

Zoltan Kunszt
Eidg. Technische Hochschule
Zurich, Switzerland

Davison Soper
University of Oregon
Eugene, Oregon 97403

December 1990

A sensible jet definition should be:

1) Simple to implement in experimental analysis
2) Simple to implement in theoretical calculations
3) Defined at any order of perturbation theory
4) Yields finite cross section at any order of perturbation theory
5) Yields a cross section that is relatively insensitive to hadronization

Purpose: “undo” parton evolution to define the “hard scattering” process

A look back: Snowmass accord 1990 ...

For theory:

- Infrared collinear (IRC) safety
- Small sensitivity of 'inclusive observables' to parton-shower & hadronisation

A sensible jet definition should be:

1) Simple to implement in experimental analysis
2) Simple to implement in theoretical calculations
3) Defined at any order of perturbation theory
4) Yields finite cross section at any order of perturbation theory
5) Yields a cross section that is relatively insensitive to hadronisation

Purpose: "undo" parton evolution to define the "hard scattering" process

Jets at the LHC

Many proposals of jet algorithms since ‘90:

- Cone-based algorithms: PxCone, midpoint, seedless, SISCone, …
- 2-to-1 recombination algorithms: C/A, Jade, kT, anti-kT, …

The **standard** algorithm for the LHC is

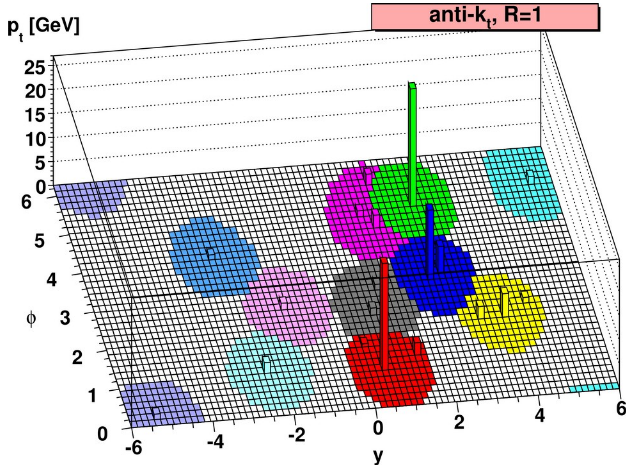
the **anti-kT algorithm**:

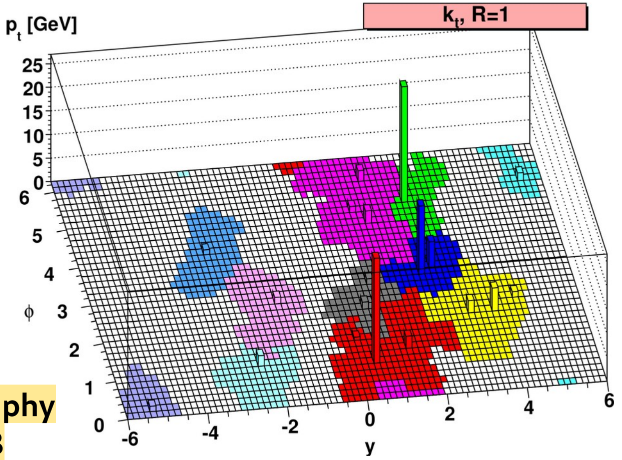
$$d_{ij} = \min(k_{T,i}^{-2}, k_{T,j}^{-2}) R_{ij}^2 \qquad d_i = k_{T,i}^{-2}$$

→ nice geometric properties

→ theoretically okay

→ insensitive to soft physics, pile up, etc.



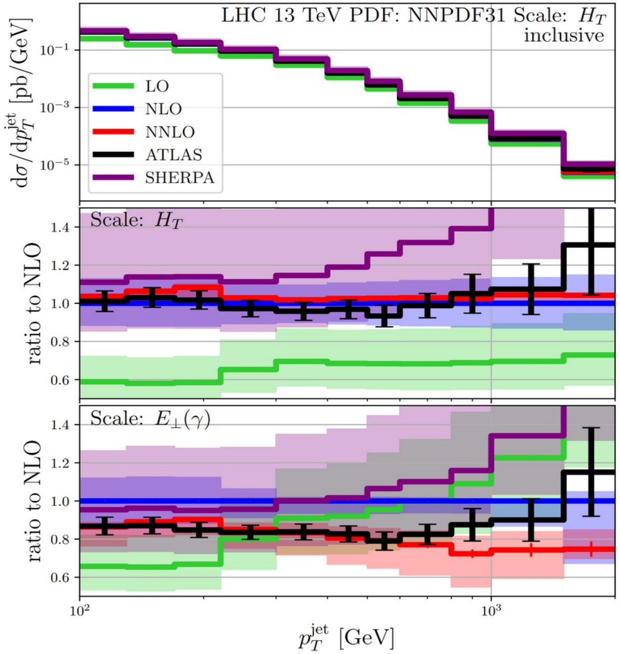


Towards Jetography
Salam 0906.1833

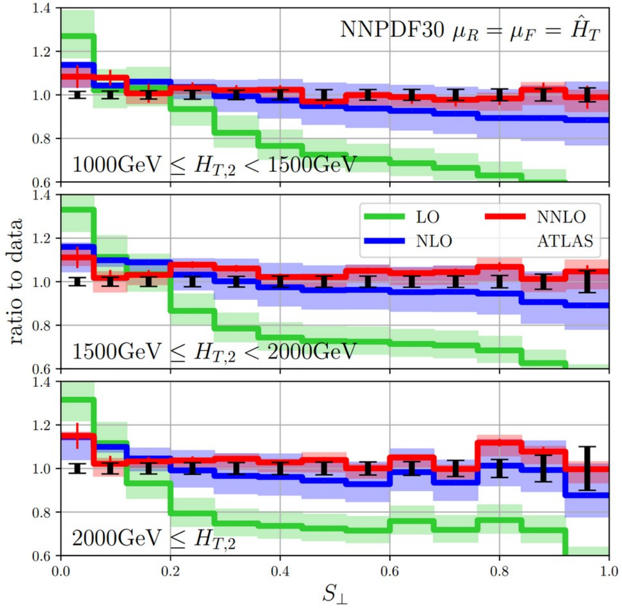
Precision comparisons of jet cross sections

Following these guidelines means that we can compare theory and experiment even though theorist talk about quarks+gluons and experimentalists about particles

$pp \to \gamma jj$ [2304.06682]

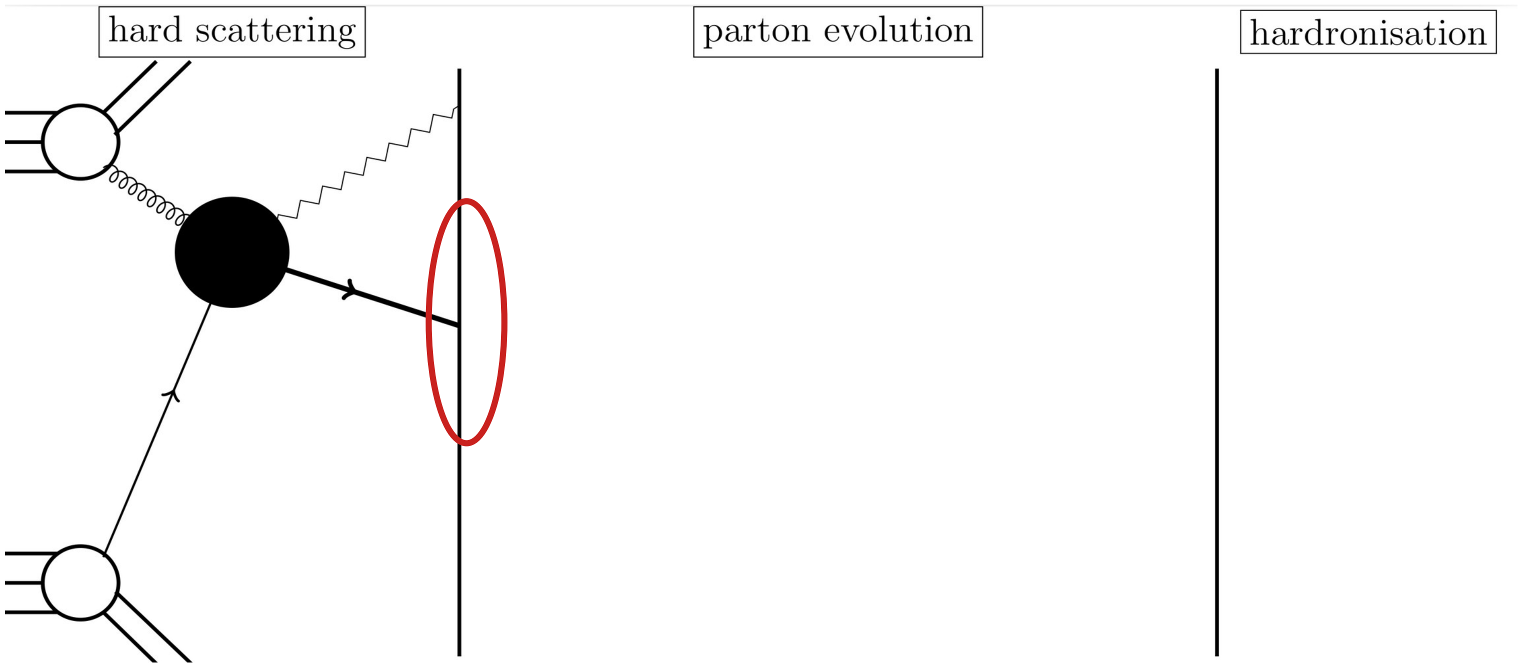


$pp \to jjj$ [2301.01086]



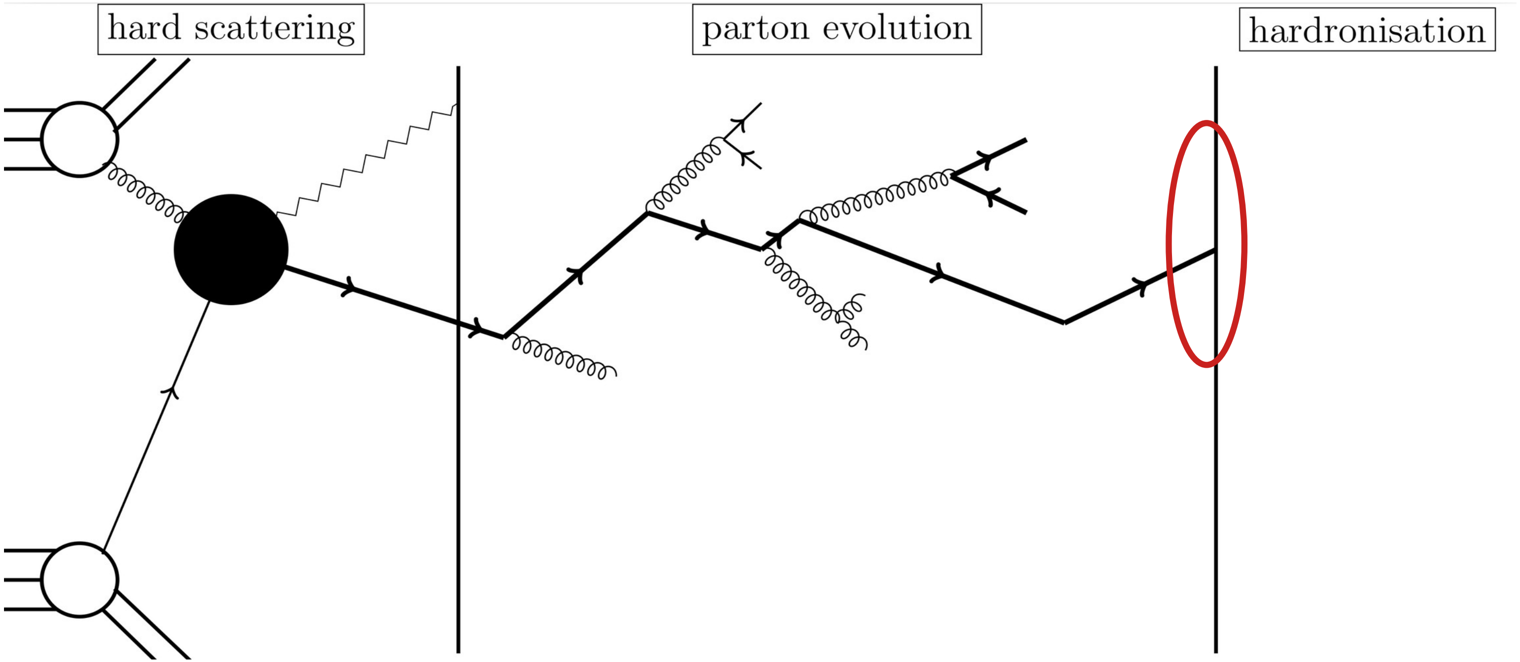
Flavoured jets

jet that initiated from a "hard scatter" product of specific flavour:
bottom, charm , "quark/gluon"



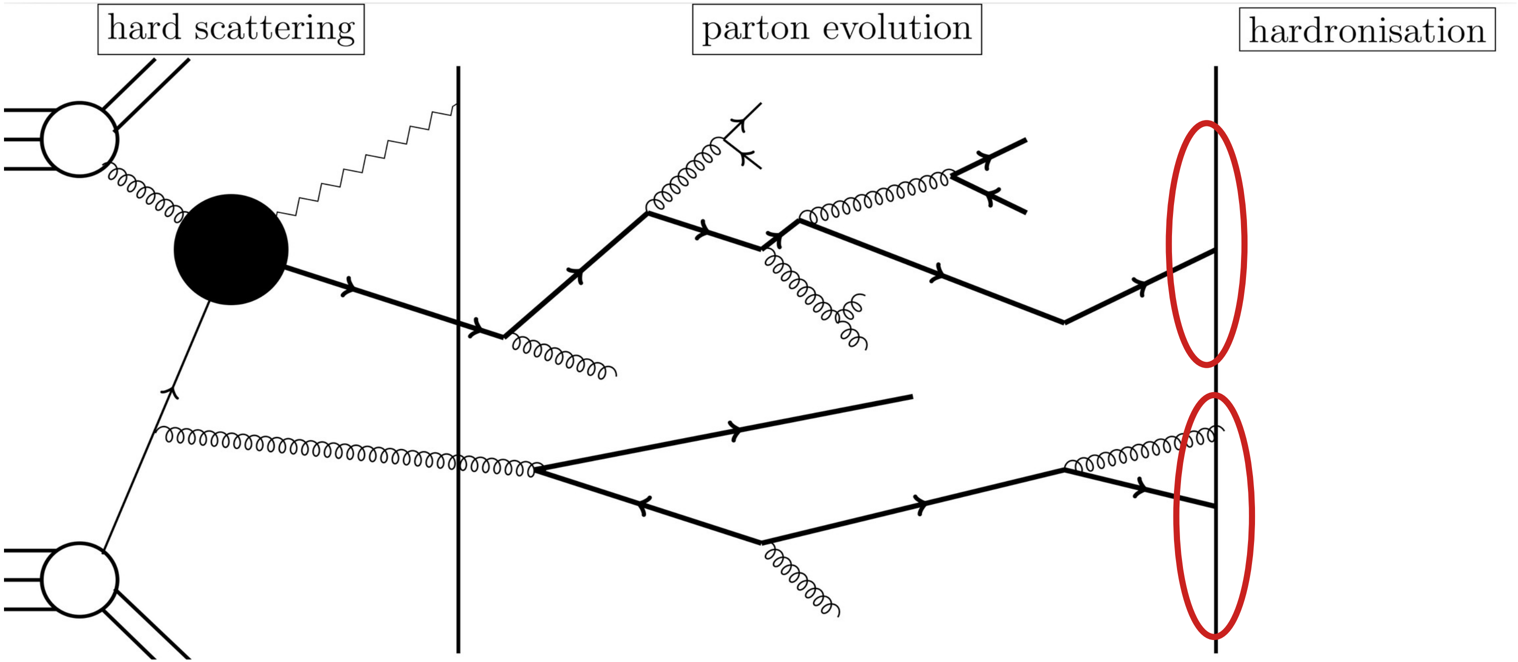
Flavoured jets

jet that initiated from a "hard scatter" product of specific flavour:
bottom, charm , "quark/gluon"



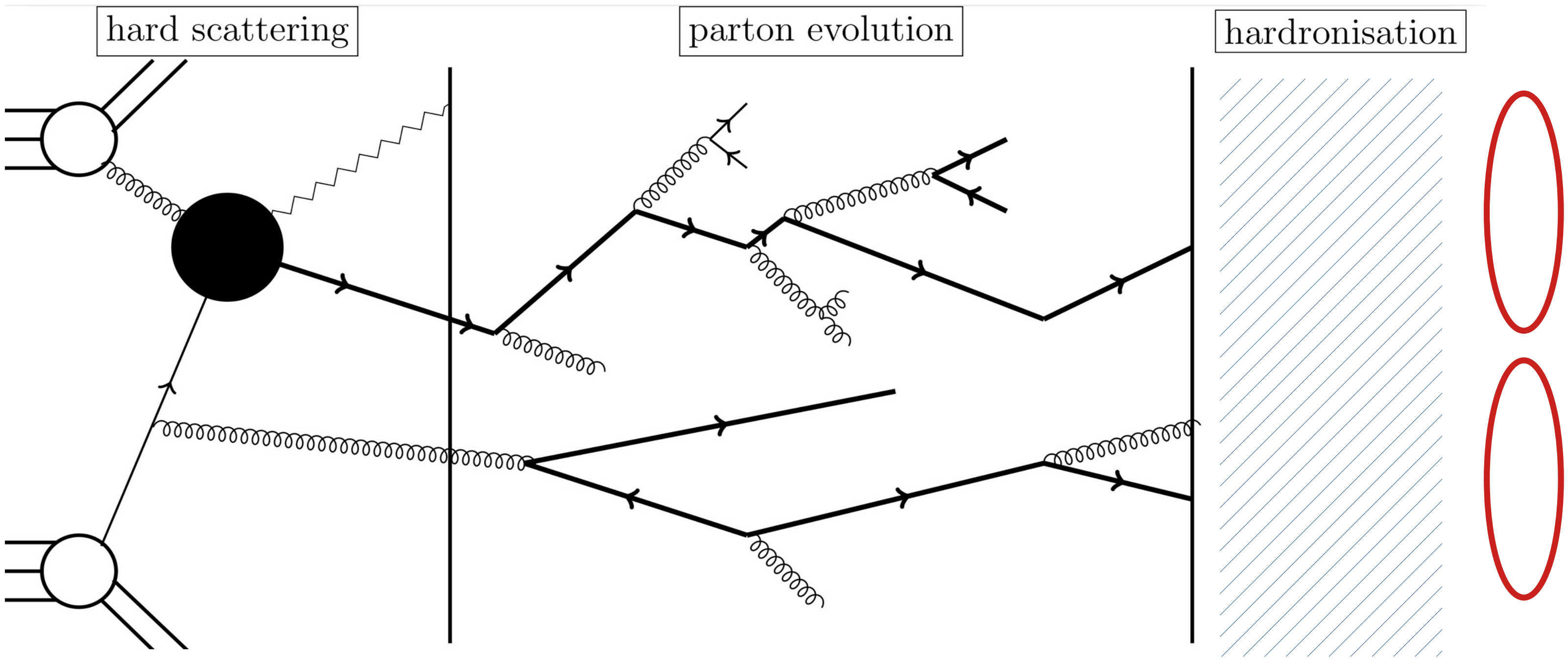
Flavoured jets

jet that initiated from a "hard scatter" product of specific flavour:
bottom, charm , "quark/gluon"



Flavoured jets

jet that initiated from a "hard scatter” product of specific flavour:
bottom, charm , "quark/gluon"



Example for experimental 'truth level' flavour tagging

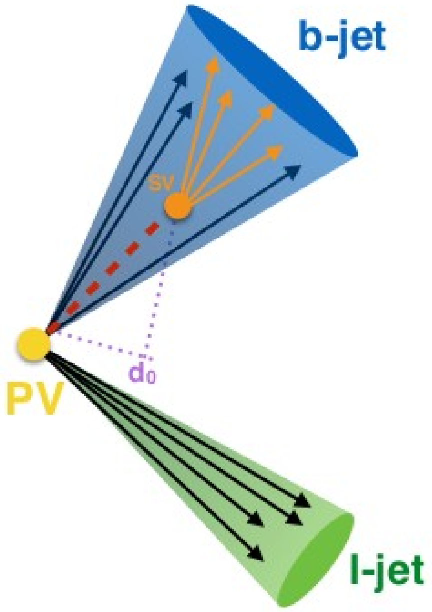
Example definition for experimental tagging

A 'truth-level' jet is defined as flavoured if:

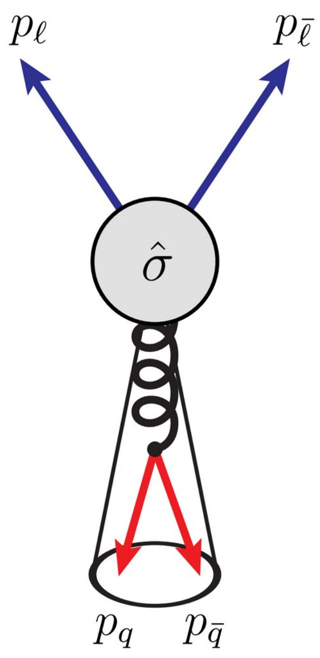
1) it contains at least one B hadron
   FO: IRC-unsafe because of $g \to b\,\bar{b}$ splitting

2) with pT > pT_cut
   FO: collinear unsafe $b \to b\,g$ splitting
   (okay in fragmentation approach)

3) within dR < R of jet axis
   FO: IRC-unsafe because soft wide angle emission



Infrared safety issues with flavoured jets I



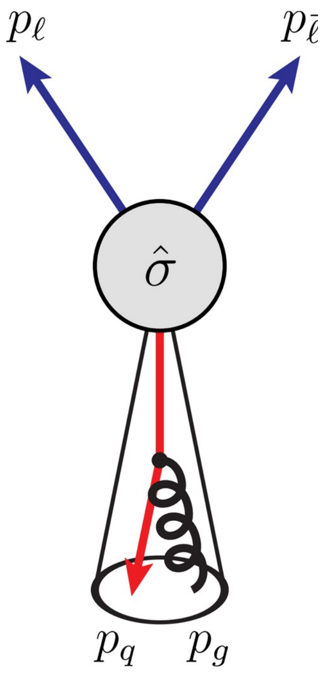
Picture from
[Gauld et al. 2302.12844]

- IRC unsafe due to g→quark-anti-quark splitting
  - Quarks massless: cross-section not defined
  - Quarks massive: logarithmic sensitivity to quark mass
- Can resolved by proper flavour recombination schemes:

scheme \ jet contents	b	$b+\bar{b}$	$b+b$	
"any flavour"	b	b	b	simplest experimentally (but collinear unsafe for $m_b \to 0$)
net flavour	b	g	$2b$	theoretically "ideal" definition; but not robust wrt B–Bbar oscillations
flavour modulo 2	b	g	g	theoretically OK; robust wrt B–Bbar oscillations

[Salam]

Infrared safety issues with flavoured jets II

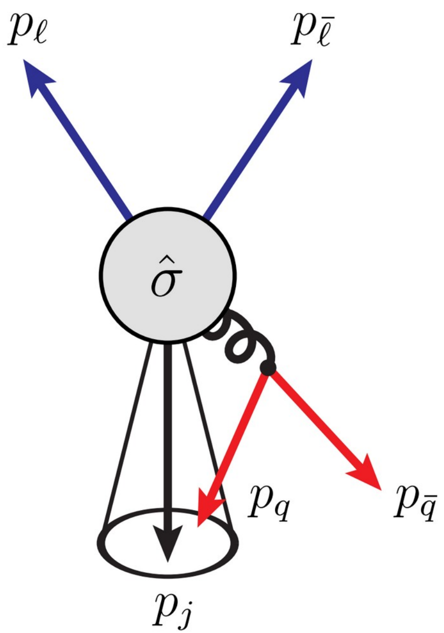


Picture from
[Gauld et al. 2302.12844]

- Collinear unsafe if pT requirement on the quark is present
- Not implementable in pQCD with massless quarks
  → proper treatment needs fragmentation functions
  → NNLO QCD example:

  B-hadron production in NNLO QCD: application to LHC ttbar events with leptonic decays, Czakon, Generet, Mitov and Poncelet, 2102.08267

Infrared safety issues with flavoured jets III



Picture from
[Gauld et al. 2302.12844]

- Starting at NNLO QCD:
  → Soft singularity from quark pairs
  - Massless quarks → cross section not defined
  - Massive quarks → logarithmic IRC sensitivity $\ln\left(\frac{m}{p_T}\right)$
- Needs modified jet algorithms!

Solution: Modified jet algorithms

These issues are known since 2006… a solution as well:

Standard kT algorithm:

Pair distance:

$$d_{ij} = \min(k_{T,i}^2, k_{T,j}^2) R_{ij}^2$$

$$R_{ij}^2 = (\Delta\phi_{ij}^2 + \Delta\eta_{ij}^2)/R^2$$

"Beam" distance for determination condition:

$$d_i = k_{T,i}^2$$

Flavour kT algorithm:

Infrared safe definition of jet flavor, Banfi, Salam, Zanderighi hep-ph/0601139

Pair distance:

$$d_{ij} = R_{ij}^2 \begin{cases} \max(k_{T,i}, k_{T,j})^\alpha \min(k_{T,i}, k_{T,j})^{2-\alpha} & \text{softer of i,j is flavoured} \\ \min(k_{T,i}, k_{T,j})^\alpha & \text{else} \end{cases}$$

Beam distance:

$$d_{i,B} = \begin{cases} \max(k_{T,i}, k_{T,B}(y_i))^\alpha \min(k_{T,i}, k_{T,B}(y_i))^{2-\alpha} & \text{i is flavoured} \\ \min(k_{T,i}, k_{T,B}(y_i))^\alpha & \text{else} \end{cases}$$

$$d_B(\eta) = \sum_i k_{T,i}(\theta(\eta_i - \eta) + \theta(\eta - \eta_i) e^{\eta_i - \eta}$$

$$d_{\bar{B}}(\eta) = \sum_i k_{T,i}(\theta(\eta - \eta_i) + \theta(\eta_i - \eta) e^{\eta - \eta_i}$$

→ problem for LHC: this is a kT algorithm → 'apples to apples' comparison not possible

New proposals for flavour-safe anti-kT jets

- Flavour with Soft-drop
  - **Practical Jet Flavour Through NNLO**
    Caletti, Larkoski, Marzani, Reichelt 2205.01109
- Flavour anti-kT
  - **Infrared-safe flavoured anti-kT jets**,
    Czakon, Mitov, Poncelet 2205.11879
- Fragmentation approach
  - **A Fragmentation Approach to Jet Flavor**
    Caletti, Larkoski, Marzani, Reichelt 2205.01117
  - **B-hadron production in NNLO QCD: application to LHC ttbar events with leptonic decays**,
    Czakon, Generet, Mitov and Poncelet, 2102.08267
- Flavour dressing → standard anti-kT + flavour assignment
  - **QCD-aware partonic jet clustering for truth-jet flavour labelling**
    Buckley, Pollard 1507.00508
  - **A dress of flavour to suit any jet**
    Gauld, Huss, Stagnitto 2208.11138
- Interleaved flavour neutralisation
  - **Flavoured jets with exact anti-kT kinematics and tests of infrared and collinear safety**
    Caola, Grabarczyk, Hutt, Salam, Scyboz, Thaler 2306.07314
- TBC...

New proposals for flavour-safe anti-kT jets

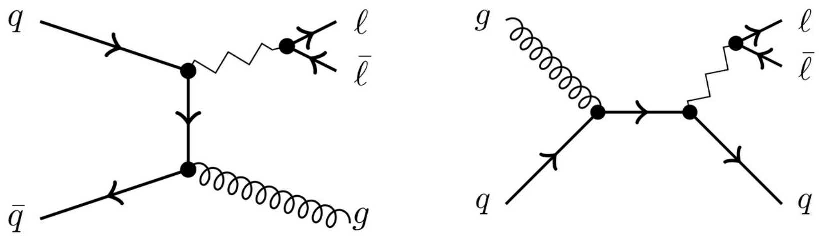
IRC-safe anti-kT flavoured jet algorithm? Yes, but which one?

Criteria:

- IRC-safety ← Highly desirable: fulfilled by all candidates (at least through NNLO)
- "truthfully" reconstruct reconstruct the original "hard" object
  → insensitive to PS+HAD+SOFT ← Desirable: robust theory predictions!
- experimentally implementable ← comment at the end
- numerically efficient ← not yet the focus of the effort
  → important for experimental implementation
- easy to implement in analysis ← wip towards full release
  → FastJet-contrib (test implementations: https://github.com/jetflav/)
- Jet-substructure?
- ...

Les Houches "FlavourFest"

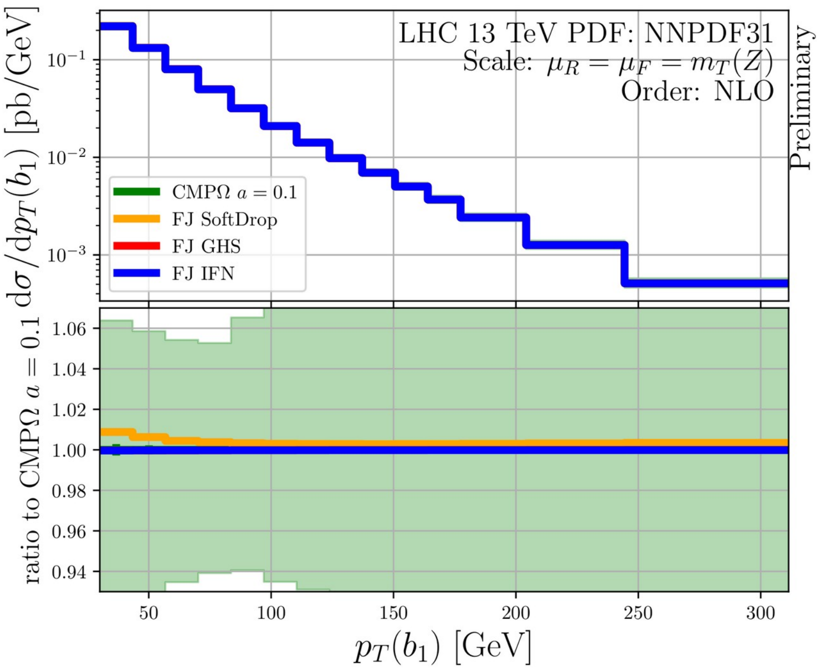
$$pp \to Z + \text{jet}/b - \text{jet}/c - \text{jet}$$

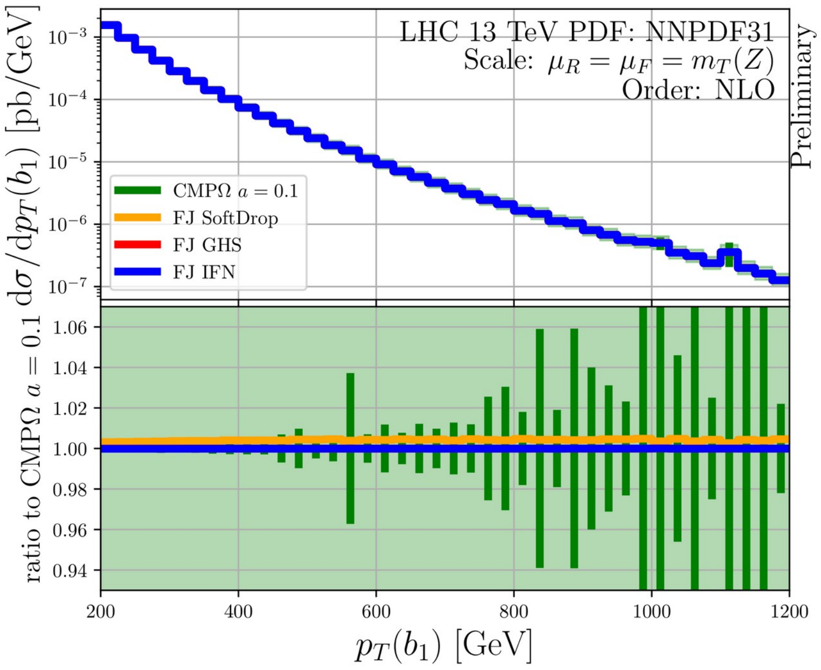


- @13 TeV
- Algorithms: Flavour anti-kT (**CMP**), Flavour dressing (**GHS**), Interleaved flavour neutralisation (**IFN**), Soft-drop (**SDF**)
- pQCD computations: up to NNLO QCD in the nf=5 scheme → massless b, c quarks
- NLO PS matched calculations:
  - SHERPA (massive quarks, dipole)
  - HERWIG7 (massive quarks, angular) and HERWIG7 (massless quarks, dipole)
- Parton-level and Particle-level

Many people contribute to this: Simone Marzani, Arnd Behring, Daniel Reichelt, James Whitehead, Andrzej Siódmok, Ludovic Scyboz, Gavin Salam, Ezra Lesser, Giovanni Stagnitto, Rene Poncelet, …

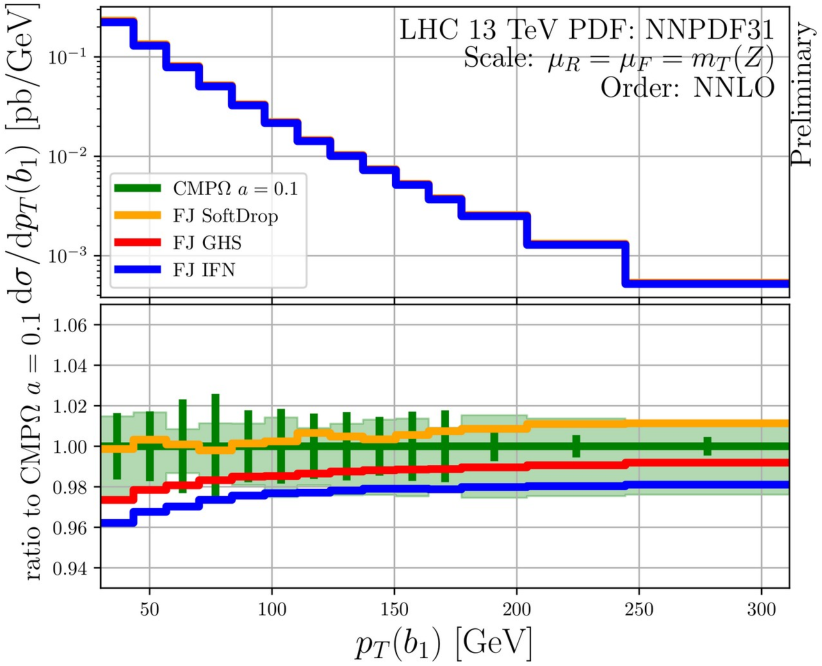
Fixed-order comparisons NLO QCD

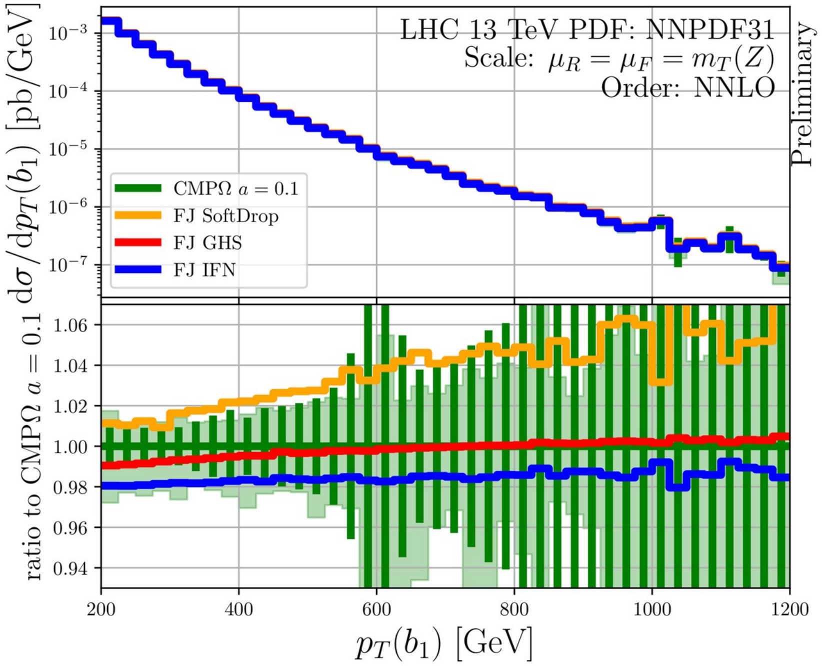




- minimal differences at FO 2-5%
- overall consistent definition of the "hard" object
- other processes tested: WH, W+charm, ttbar+decays

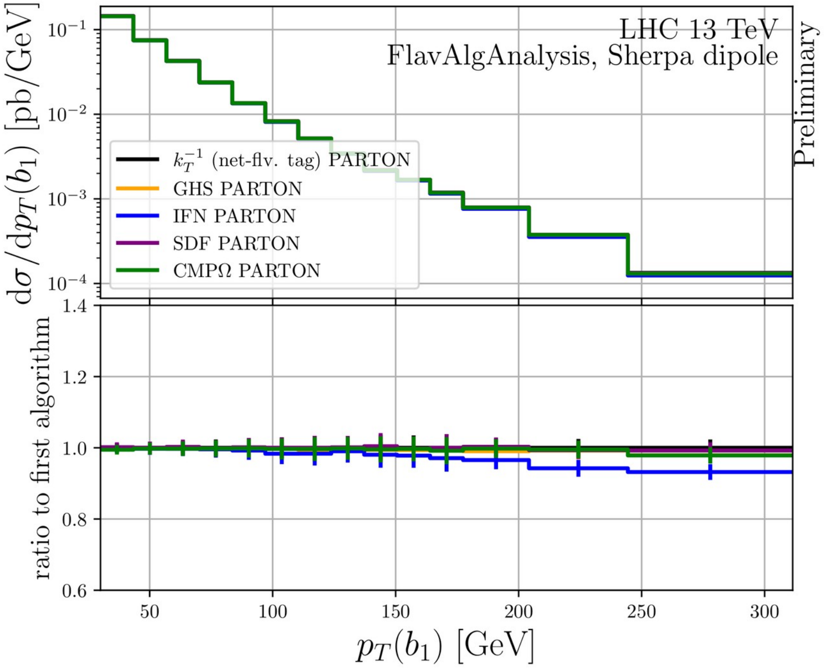
Fixed-order comparisons NNLO QCD

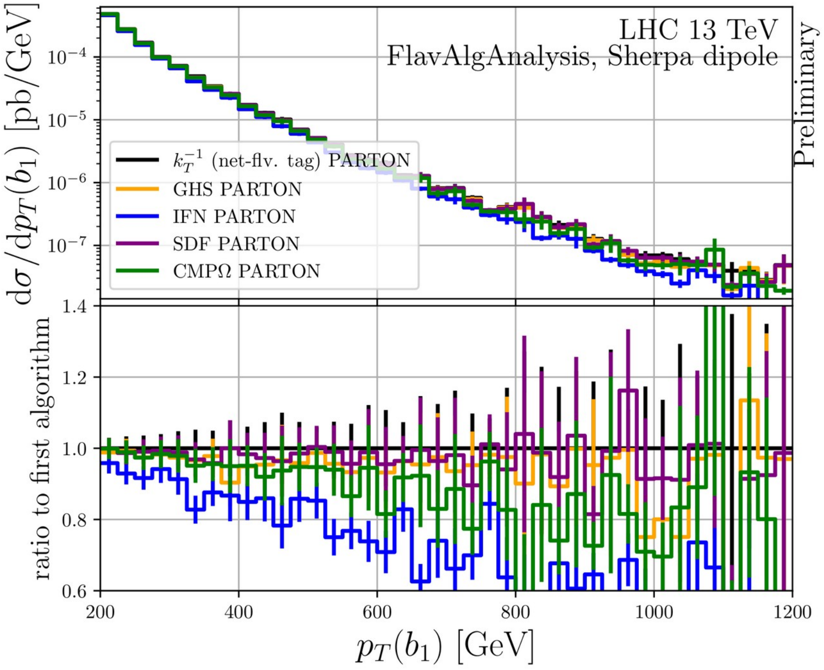




- minimal differences at FO 2-5%
- overall consistent definition of the "hard" object
- other processes tested: WH, W+charm, ttbar+decays

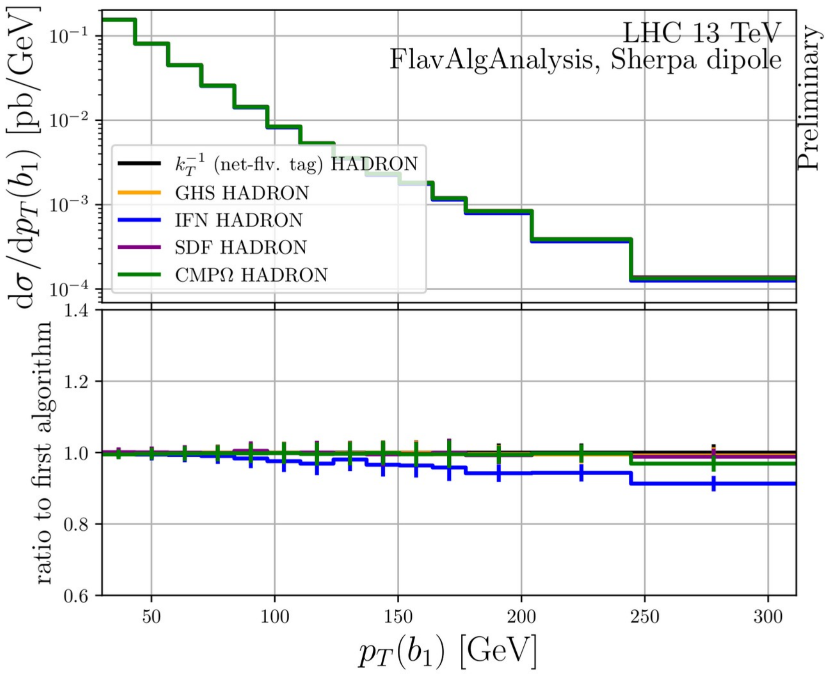
NLO+PS at parton-level

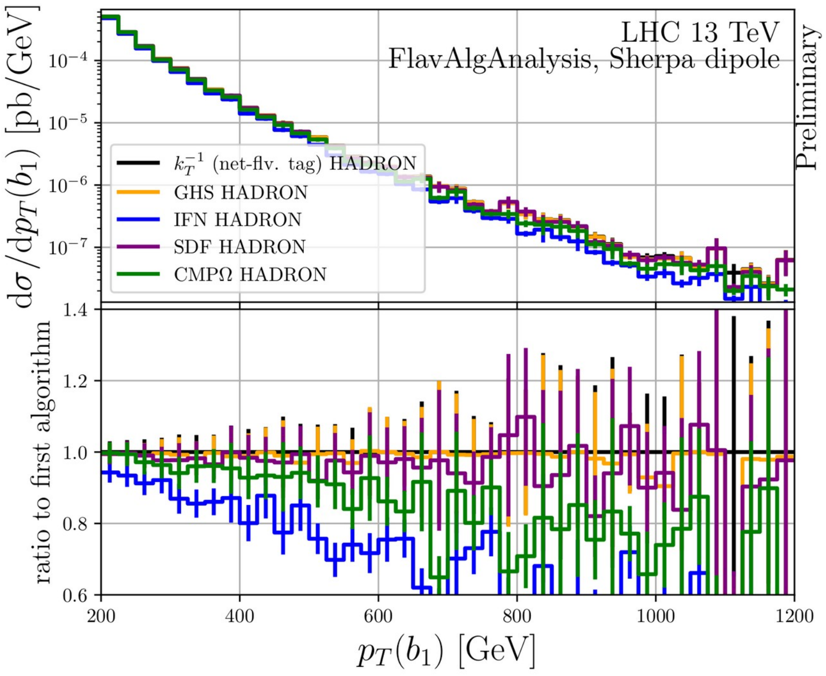




Larger differences → in particular in high pT tail
→ IFN removes flavour more aggressive than CMP > GHS > SDF

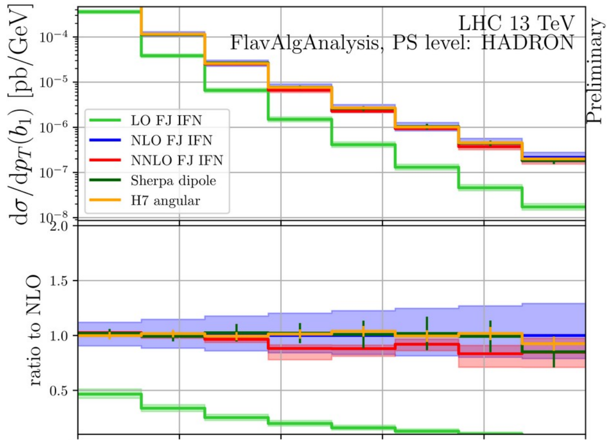
NLO+PS at particle-level

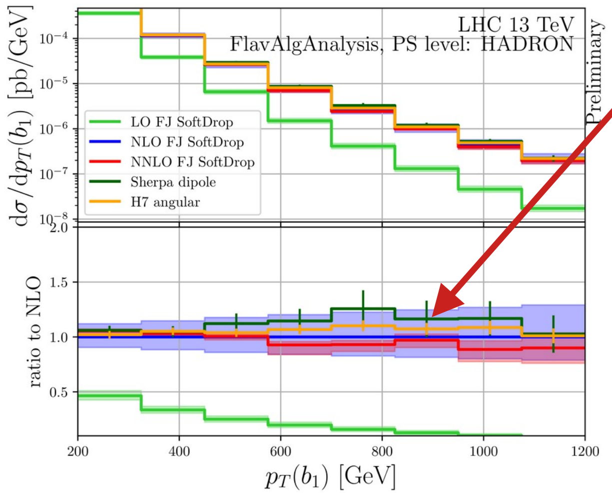




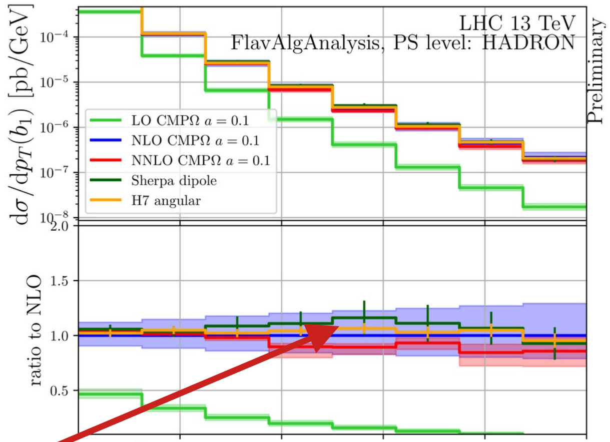
Larger differences → in particular in high pT tail
→ observations insensitive to hadronisation effects

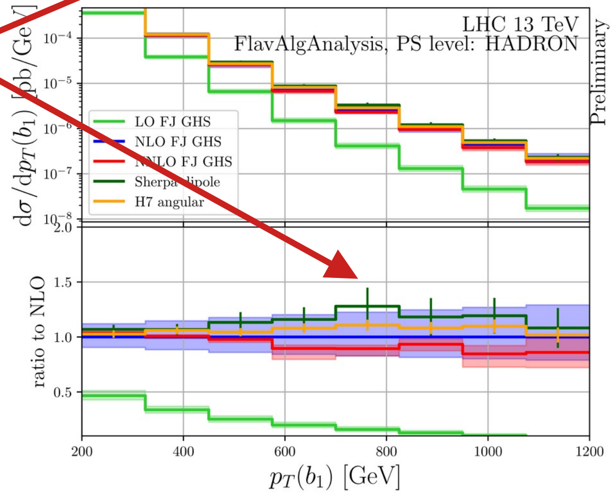
Fixed-order and NLO+PS comparisons: b-jets



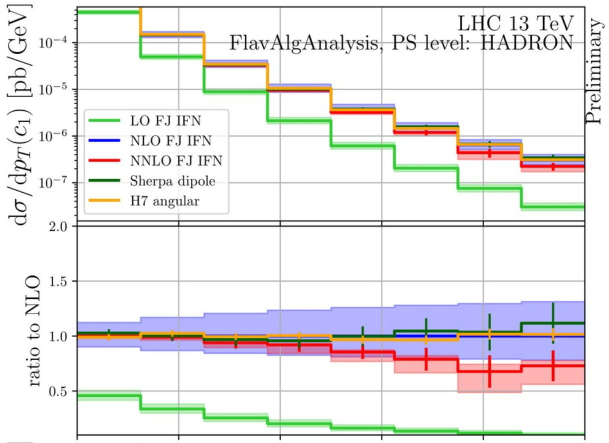


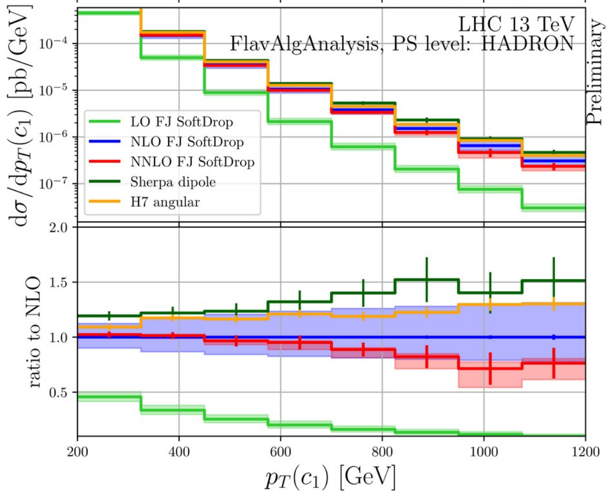
- Overall good agreement
- SDF, CMP & GHS show some shape & PS model dependence
- IFN more stable
- H7 and SHERPA give consistent results





Fixed-order and NLO+PS comparisons: c-jets

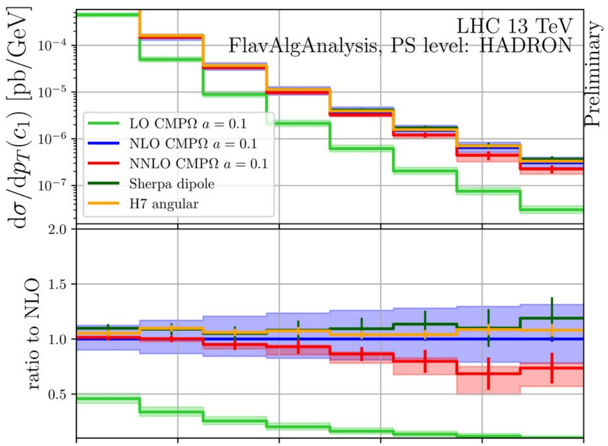


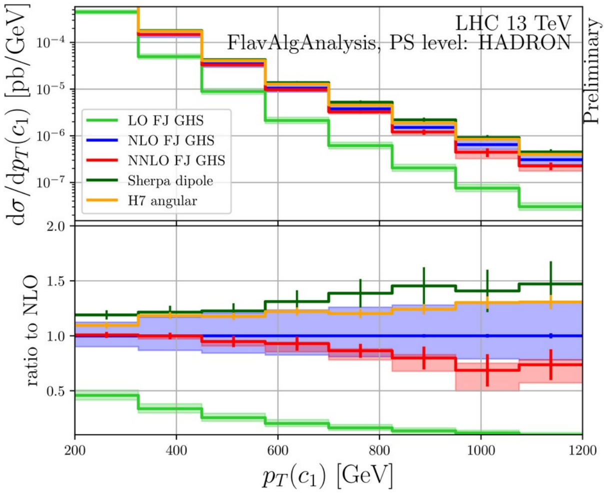


Similar to b-jets but differences enhanced

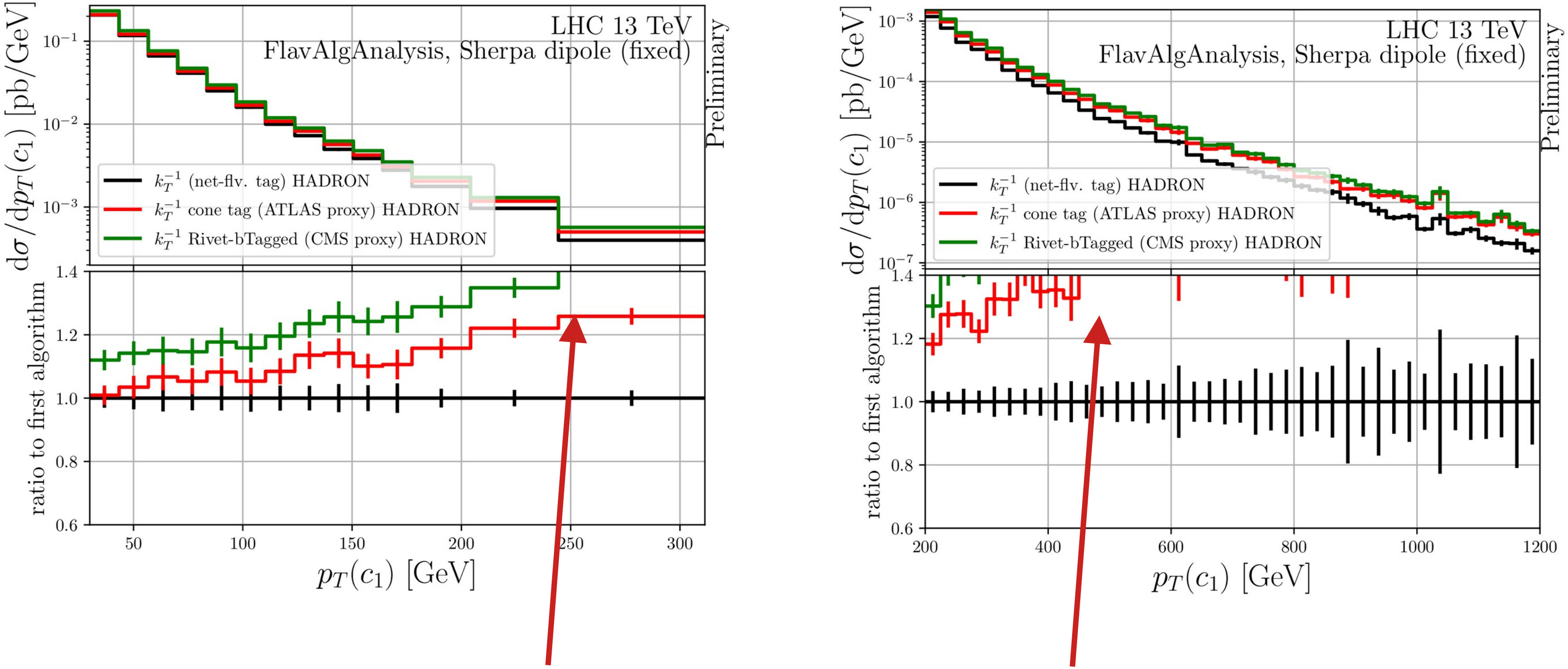
→ smaller mass leads to larger flavour abundance

→ enhances sensitivity to flavour treatment





Comparison of experimental tagging



Net-flavour tagging makes already a huge difference!
Driven by g → cc splittings

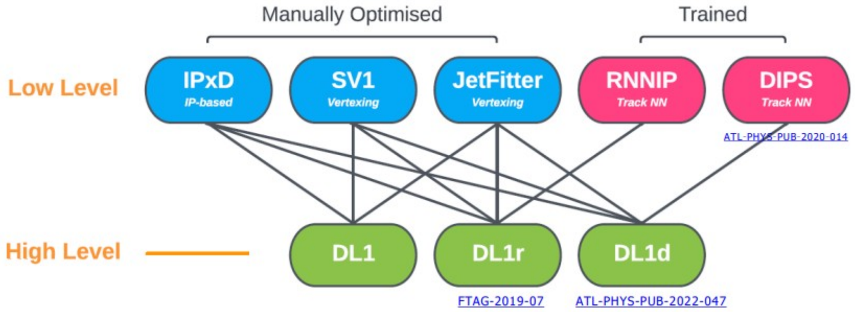
Towards experimental implementation

- The flavoured jet algorithms require detailed flavour information
  → flavour algorithms difficult to implement experimentally
  Limited by detector-resolution & efficiencies!

1) Unfolding (that is done so far):
   - g → b b splitting if both b's hadronise to B-hadrons
     (this is different to b $\bar{b}$ = g @ fixed order)
   - Hadronisation/non-perturbative models
   - Unfolding corrections can be sizeable O(5-10%) and relies on IR sensitive anti-kT
2) Improvement on experimental side:
   → Potential improvements if g → bb splittings can be captured experimentally
3) Using IRC-safe truth labels in ML – tagger training

Experimental b/c-tagging with NN

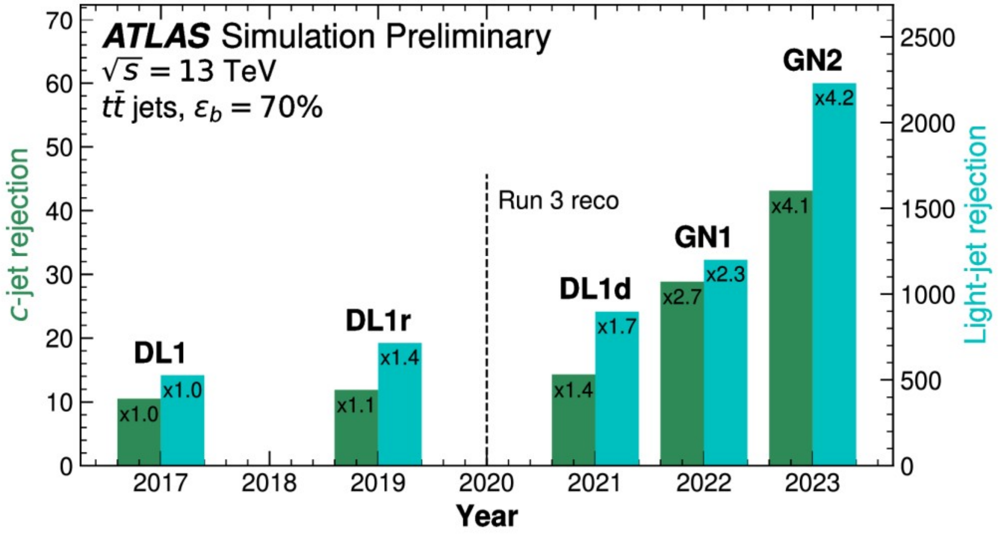
Credit: Arnaud Duperrin (DIS23 talk)



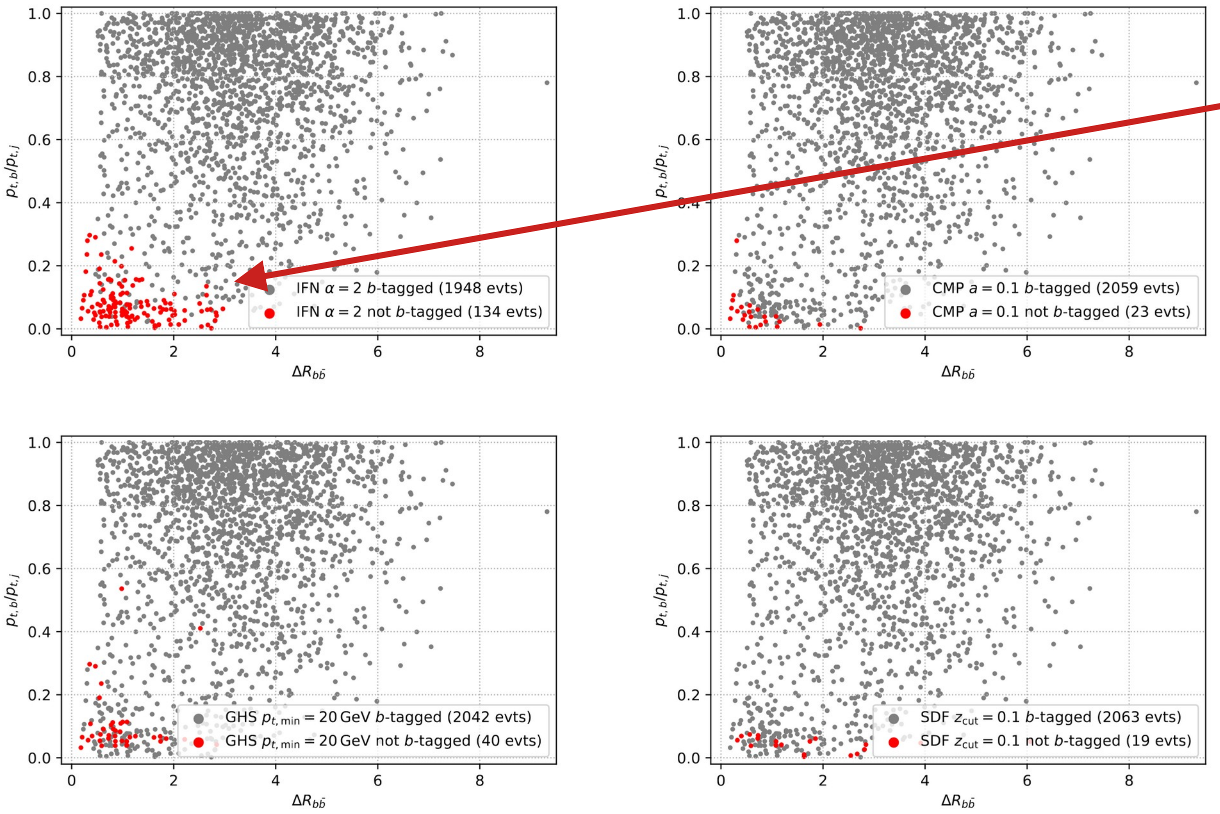
Using NN to perform b-tagging

- Many Run II/III analysis use already NN based taggers
- For example ATLAS: DL1
  → uses precomputed low-level infos
- Next generation will directly use hit, track and jet information
  → further performance boost

The truth level information comes (partially or indirectly) from MC simulations

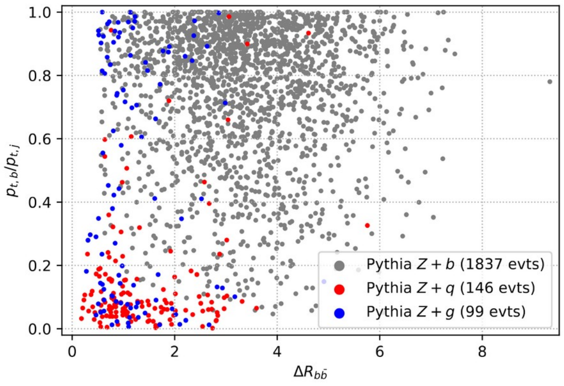


Truth-level input



Red dots show jets that would be removed w.r.t. to anti-kT

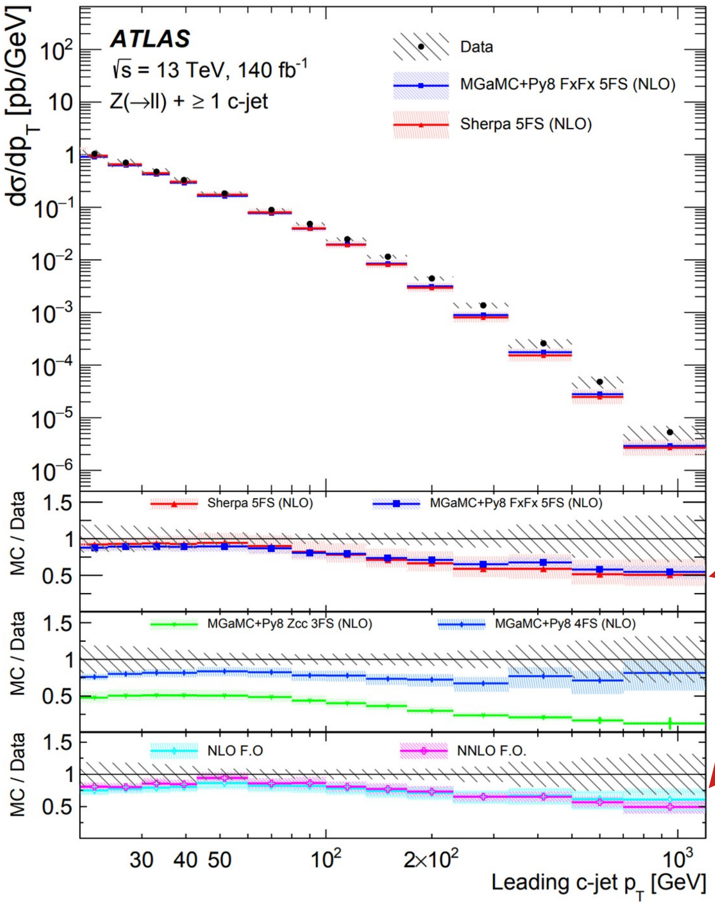
Red dots show 'light' jets



[Thanks to Ludovic Scyboz & Gavin Salam]

Take away:
Flavoured algorithms remove wrongly tagged jets!

Example flavoured jet measurement



Measurements of the production cross-section for a Z boson in association with b- or c-jets in proton-proton collisions at sqrt{s}=13 TeV with the ATLAS detector, 2403.15093

- Using unfolding to compare to GHS algorithm

Clear mismodelling of high-pT tails
→ likely to due g→bb splittings

Take home messages

- Accurate modelling of (heavy) flavour jets requires improvement on the jet definition →needed for precision phenomenology
- New flavoured jet-algorithms provide IRC safe definitions
- Les Houches Study to study qualitative and quantitative differences between proposals → implementation in fastJet framework
- Experimental implementation still an open questions
  - Unfolding? → Large uncertainties (still uses the IRC unsafe anti-kT jets)
  - Improvement on tagging procedures? Challenging! (maybe g→ bb tagging?)
  - Truth label for ML-tagger training? Seems a sensible way forward!

Thanks to all the contributors to the Les Houches study!

Backup

LHC precision computations with flavoured jets

Associated Higgs production + decays in b-quarks:

Associated production of a Higgs boson decaying into bottom quarks at the LHC in full NNLO QCD
Ferrera, Somogyi, Tramontano 1705.10304

NNLO QCD corrections to associated WH production and H → bbbar decay
Caola, Luisoni, Melnikov, Röntsch 1712.06954

Associated production of a Higgs boson decaying into bottom quarks and a weak vector boson decaying leptonically at NNLO in QCD
Gauld, Gehrmann-De Ridder, Glover, Huss, Majer 1907.05836

Bottom quark mass effects in associated WH production with the H → bbbar decay through NNLO QCD
Behring, Bizoń, Caola, Melnikov, Röntsch 2003.08321

VH + jet production in hadron-hadron collisions up to order \alpha_s^3 in perturbative QCD
Gauld, Gehrmann-De Ridder, Glover, Huss, Majer 2110.12992

+Partonshower:

NNLOPS accurate associated HZ production with H → bbbar decay at NLO
Astill, Bizoń, Re, Zanderighi 1804.08141

NNLOPS description of the H → bbbar decay with MiNLO
Bizoń, Re, Zanderighi 1912.09982

Next-to-next-to-leading order event generation for VH production with H → bbbar decay
Zanoli, Chiesa, Re, Wiesemann, Zanderighi 2112.04168

LHC precision computations with flavoured jets

Vector + flavoured jet(s) production:

NLO QCD predictions for Wbbbar production in association with up to three light jets at the LHC
Anger,Cordero, Ita, Sotnikov 1712.05721

Predictions for Z-Boson Production in Association with a b-jet at O(\alpha_s^3)
Gauld, Gehrmann-De Ridder, Glover, Huss, Majer 2005.03016

NNLO QCD predictions for W+c-jet production at the LHC,
Czakon, Mitov, Pellen, Poncelet 2011.01011

NNLO QCD corrections to Wbbbar production at the LHC,
Hartanto, Poncelet, Popescu, Zoia 2205.01687

A detailed investigation of W+c-jet at the LHC,
Czakon, Mitov, Pellen, Poncelet 2212.00467

Associated production of a W boson and massive bottom quarks at next-to-next-to-leading order in QCD,
Buonocore, Devoto, Kallweit, Mazzitelli, Rottoli, Savoini, 2212.04954

NNLO QCD predictions for Z-boson production in association with a charm jet within the LHCb fiducial region
Gauld, Gehrmann-De Ridder, Glover, Huss, Rodriguez Garcia, Stagnitto 2302.12844

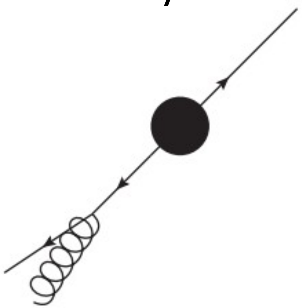
Precise QCD predictions for W-boson production in association with a charm jets
Gehrmann-De Ridder, Gehrmann, Glover, Huss, Garcia, Stagnitto, 2311.14991

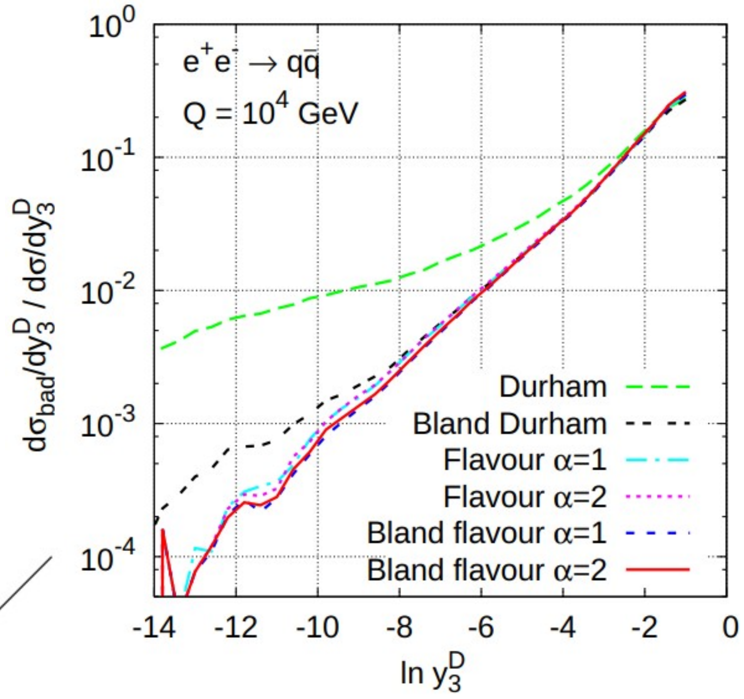
Tests of IR safety

Tests of IR safety

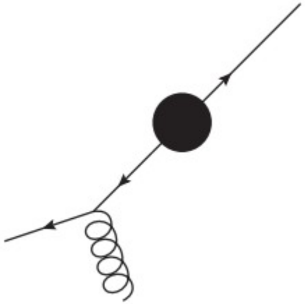
- Rate of bad-identified jet-flavour as a function of IR sensitive variable
- Parton-shower to model many emissions

Small y3



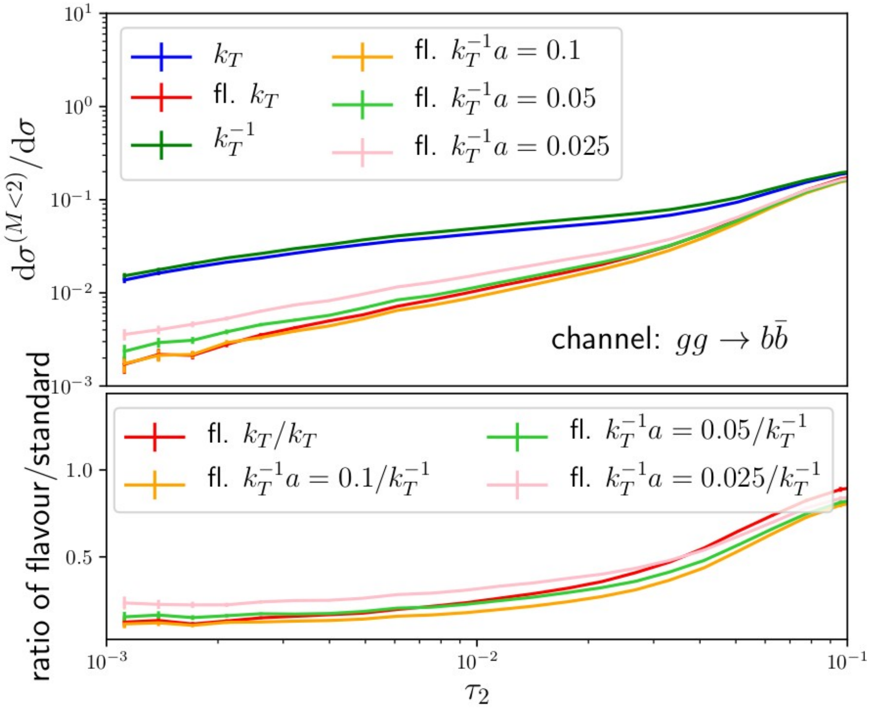


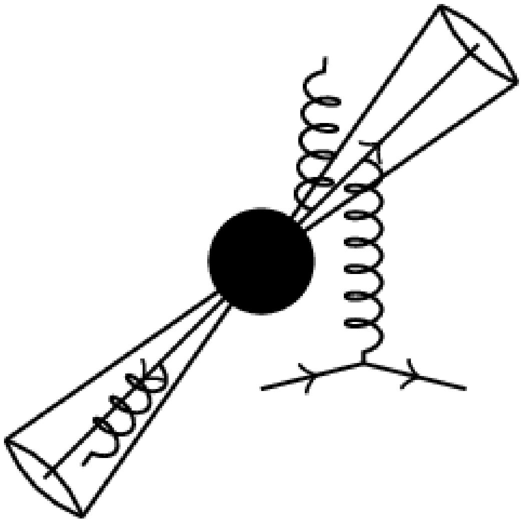
Large y3



Tests of IR safety with parton showers

In the di-jet limit the flavour needs to correspond to tree level flavours
→ misidentification rate needs to vanish in di-jet back-to-back limit
→ IR sensitive observable 2-jettiness



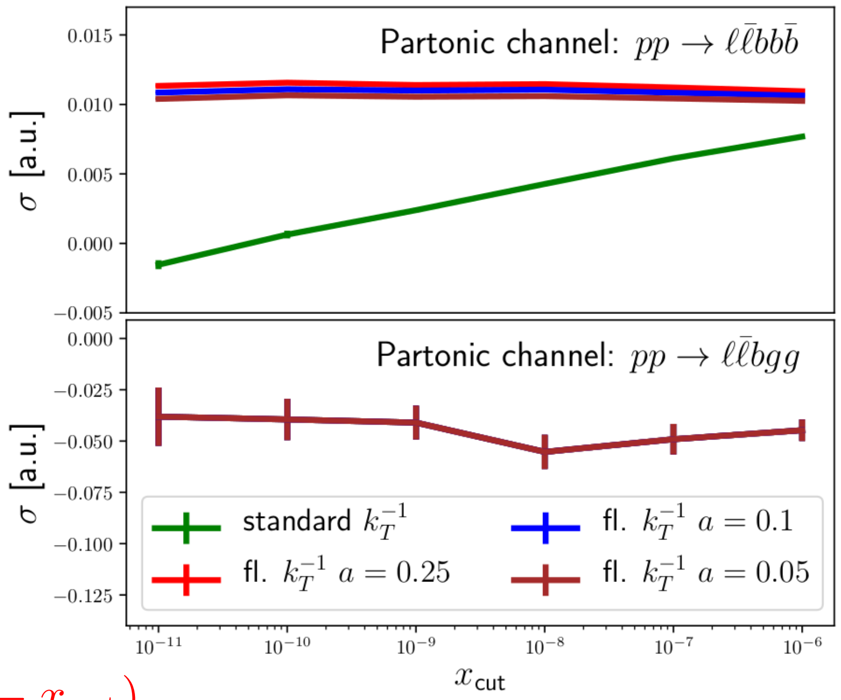


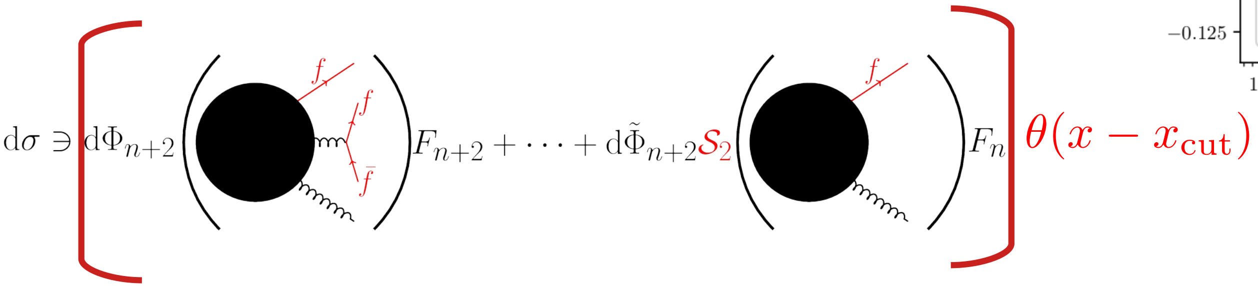
Tests of IR safety with NNLO FO computations

IR sensitivity of jet cross sections on (technical)
IR regulating parameter x

In the limit x_cut → 0:

- IR safe jet flavour → no dependence on x_cut
- IR non-safe jet flavour → logarithmic divergent



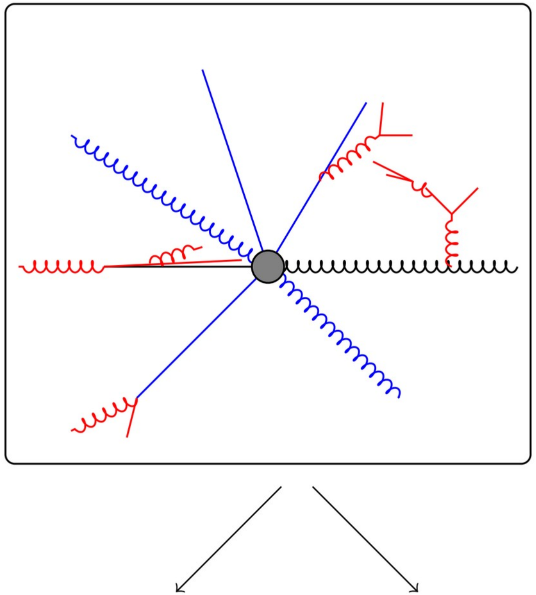


More tests...

Flavoured jets with exact anti-kT kinematics and tests of infrared and collinear safety
Caola, Grabarczyk, Hutt, Salam, Scyboz, Thaler 2306.07314

- IRC safety testing suite:

Credit: Ludo Scyboz



Set of hard jets
$\mathcal{J}_{\text{hard}} = \{(p_1, f_1), \ldots\}$

$\stackrel{!}{=}$

Set of hard+IRC jets
$\mathcal{J}_{\text{hard+IRC}} = \{(\tilde{p}_1, \tilde{f}_1), \ldots\}$

More tests...

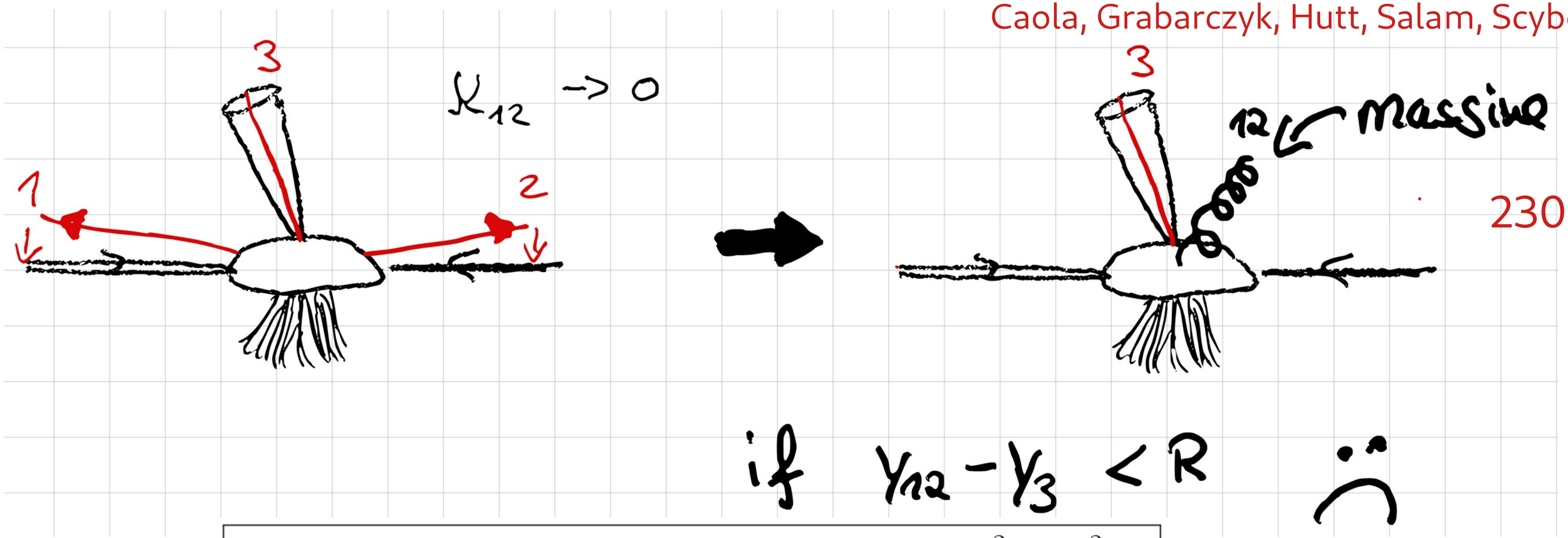
Flavoured jets with exact anti-kT kinematics and tests of infrared and collinear safety
Caola, Grabarczyk, Hutt, Salam, Scyboz, Thaler 2306.07314

order relative to Born		anti-k_t	flav-k_t ($\alpha = 2$)	CMP	$\text{GHS}_{\alpha,\beta}$ $(2,2)$	anti-k_t+IFN_α	C/A+IFN_α
α_s	FHC	✓	✓	✓	✓	✓	✓
	IHC	✓	✓	✓	✓	✓	✓
α_s^2	FDS	✗ II B	✓	✓	✓	✓	✓
	IDS	✗ II B	✓	✓	✓	✓	✓
	FHC×IHC	✓	✓	✓	✓	✓	✓
	IHC^2	✓	✓	✗ C 2	✓	✓	✓
	FHC^2	✓	✓	✓	✗ C 4	✓	✓
α_s^3	IHC×IDS		~ C 1	✗ C 3	~ C 1	✓	✓
	rest					✓	✓
α_s^4	IDS×FDS				✗ C 5	✓	✓
	rest					✓	✓
α_s^5						✓	✓
α_s^6						✓	✓

Improved distance for CMP/flavour anti-kT

Issue for double collinear limits wrt. to initial states

Many thanks to
Caola, Grabarczyk, Hutt, Salam, Scyboz,Thaler



2306.07314

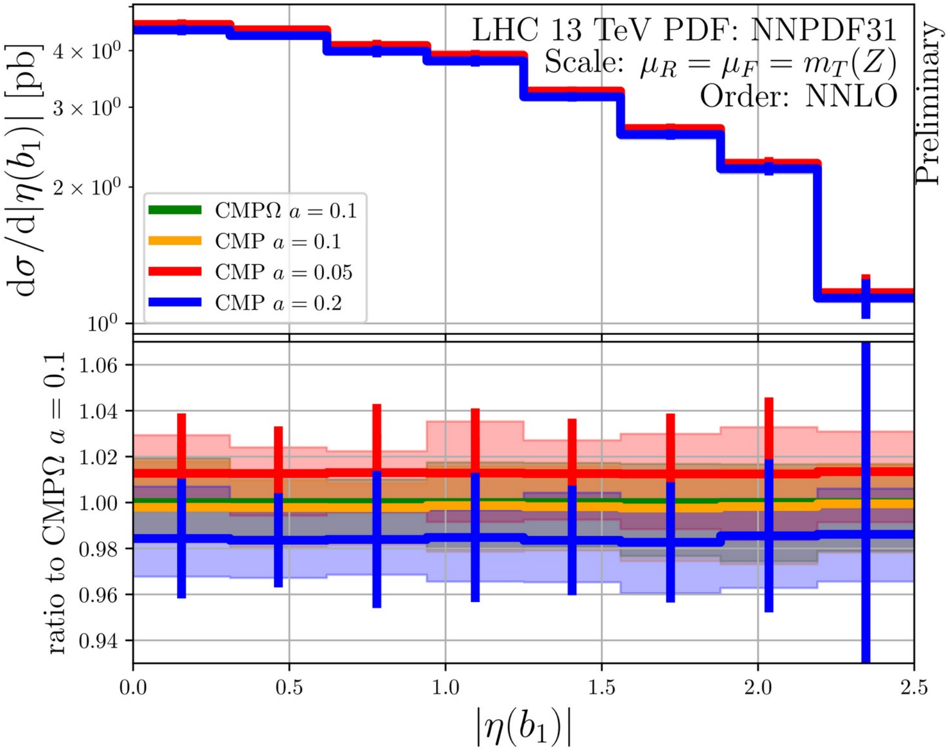
$$\mathcal{S}_{ij} \equiv 1 - \theta\left(1-\kappa_{ij}\right)\cos\left(\frac{\pi}{2}\kappa_{ij}\right) \quad \text{with} \quad \kappa_{ij} \equiv \frac{1}{a}\,\frac{k_{T,i}^2 + k_{T,j}^2}{2k_{T,\max}^2}\,.$$

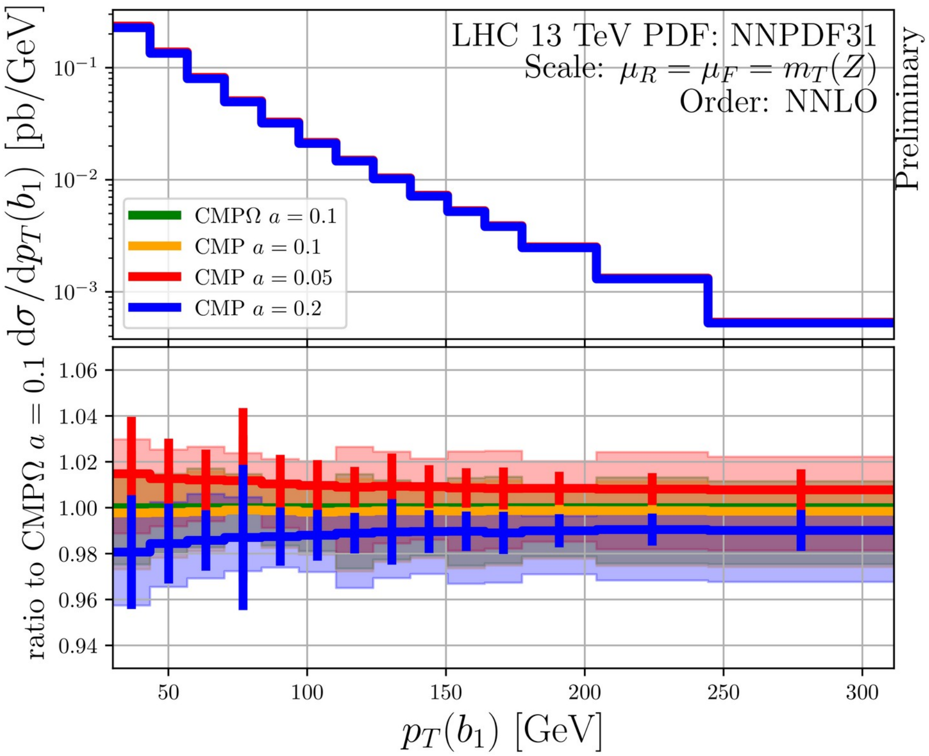
Their proposal: $\mathcal{S}_{ij} \to \overline{\mathcal{S}}_{ij} = \mathcal{S}_{ij}\,\frac{\Omega_{ij}^2}{\Delta R_{ij}^2}$ $\quad \Omega_{ik}^2 \equiv 2\left[\frac{1}{\omega^2}\left(\cosh(\omega\Delta y_{ik}) - 1\right) - \left(\cos\Delta\phi_{ik} - 1\right)\right]$

Flavour anti-kT: impact of Ω_ij

Calculations performed with sector-improved residue subtraction scheme
1408.2500 & 1907.12911

Les Houches Jet Flavour WG





Negligible difference between CMPΩ and CMP

Benchmark process: Z+b-jet

Well studied up to $\mathcal{O}(\alpha_s^3)$:

Predictions for Z-Boson Production in Association with a b-jet at O(a_s^3), Gauld, Gehrmann-De Ridder, Glover, Huss, Majer 2005.03016

- Flavour-kT algorithm
- Unfolding of experimental data (RooUnfold, bin-by-bin unfolding)
- Matching between four- and five-flavour schemes (FONLL)

$$\mathrm{d}\sigma^{\mathrm{FONLL}} = \mathrm{d}\sigma^{5\mathrm{fs}} + (\mathrm{d}\sigma^{4\mathrm{fs}}_{m_b} - \mathrm{d}\sigma^{4\mathrm{fs}}_{m_b\to 0})$$

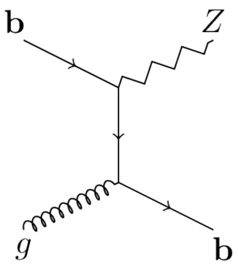
- CMS measurement @ 8 TeV

Measurements of the associated production of a Z boson and b jets in pp collisions at \sqrt{s} = 8 TeV}, CMS 1611.06507

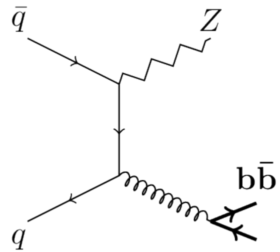
→ Ideal testing ground for flavour anti-kT

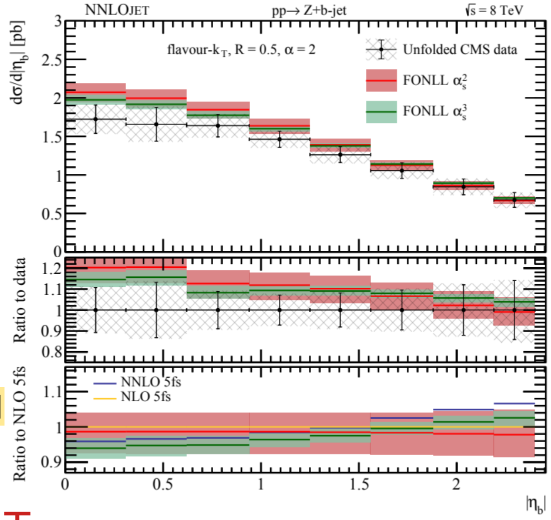
pp → Z(ll) + b-jet

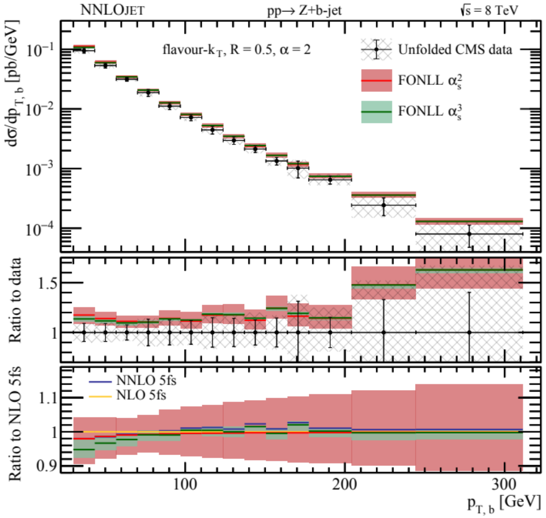
5fs:



4fs:







Bin-by-bin unfolding

Estimation of hadronisation and experimental tagging corrections
→ NLO + PS (Madraph+Pythia8)

Unfolding factor = NLO+PS (had = Off) / NLO+PS (had = On)