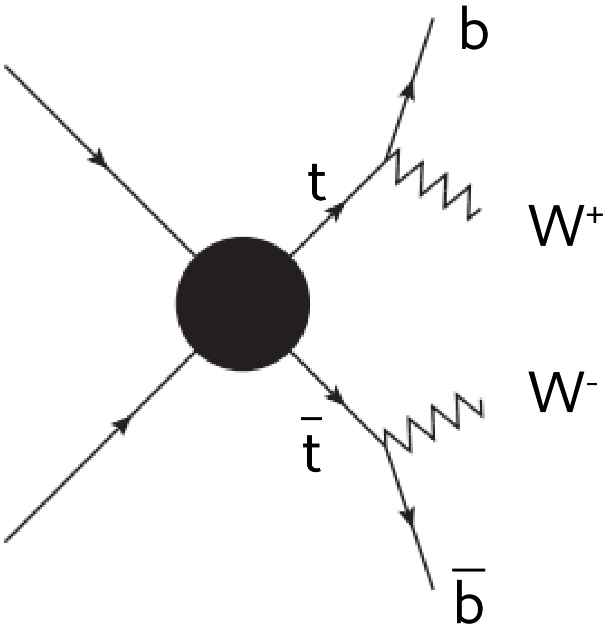
Jet identification and flavoured jet algorithms

Rene Poncelet

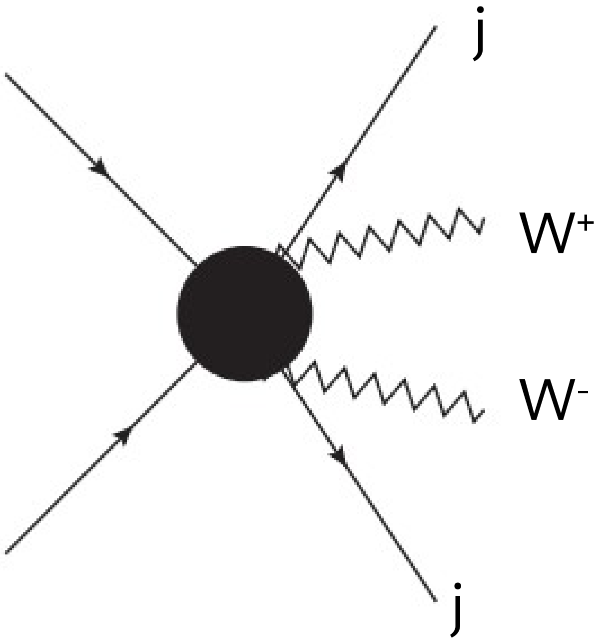
Outline

- Why are flavoured jets interesting?
- Infrared safety of flavoured jet definitions
  - What is the issue?
  - Flavoured jet algorithms
- Experimental flavour-tagging
- Phenomenology with flavour anti-kT
  - Vector-bosons + flavoured jets

Top-quark production



vs.

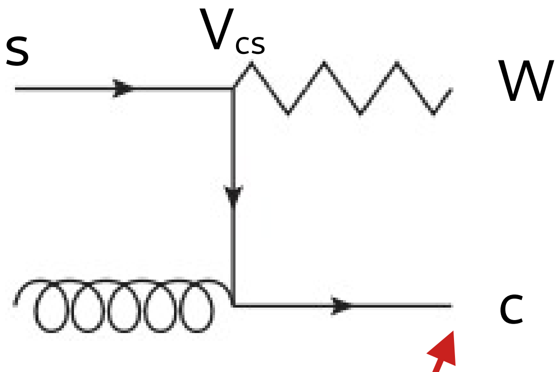


Top-quark pairs:

→ Experimental signature 2 – b-jets + WW

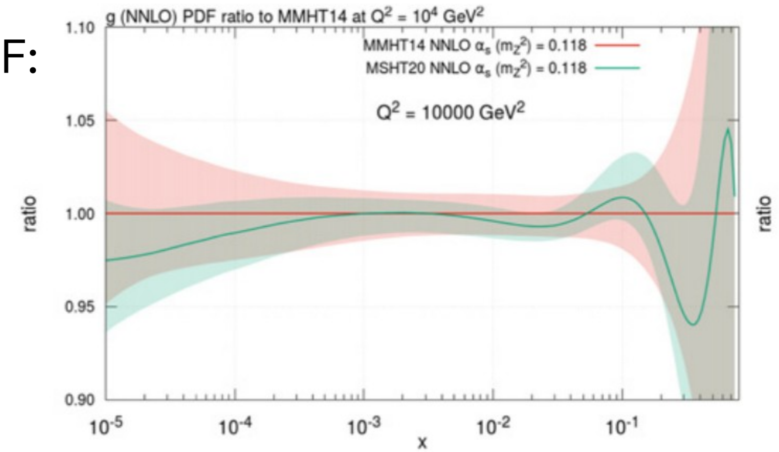
→ b-jet tagging reduces WW+QCD background dramatically.

W + charm jet

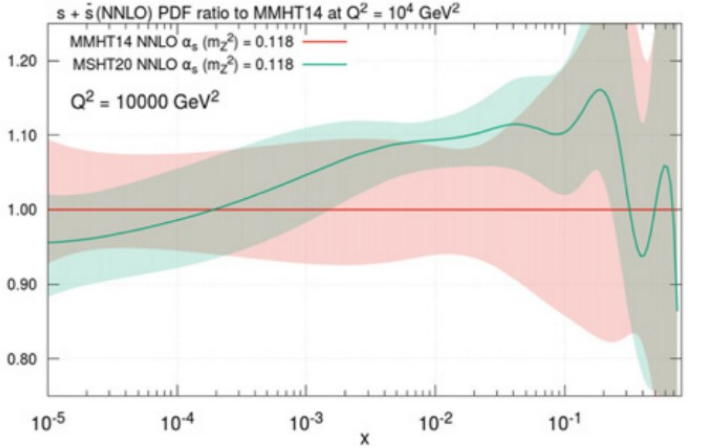


Tagging of charm jet to increase sensitivity to strange quark PDF

gluon PDF:



s+s PDF:



PDF4LHC22 [2203.05506]

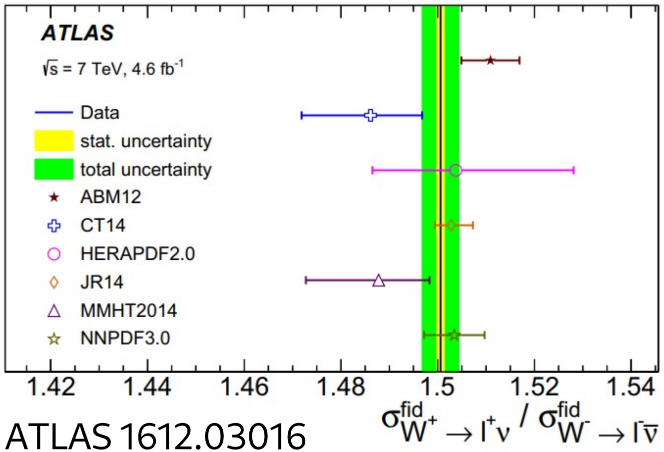
W + charm jet

Could solve long-standing puzzle:

Strange – anti – strange asymmetry

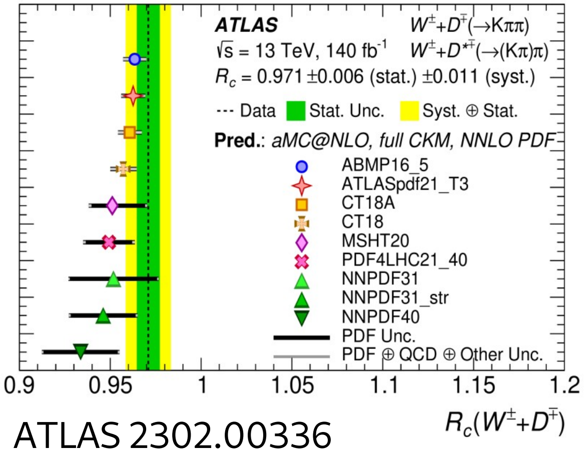
- pQCD: Three loop SM prediction q→ q’ ≠ q → $\bar{q}$’ small effect <x(s-$\bar{s}$)> ~ 10^{-4}
- Size of non-perturbative effect unknown

7 TeV analysis favours s ≠ $\bar{s}$



ATLAS 1612.03016

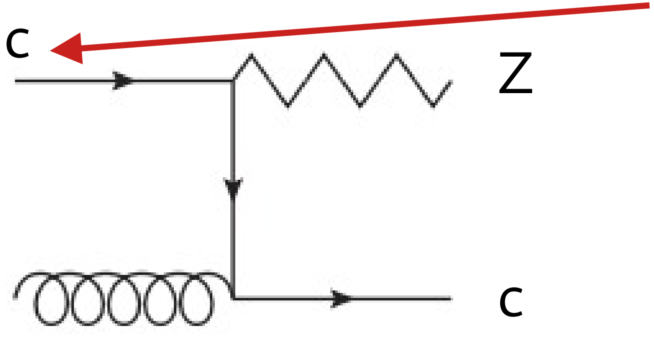
13 TeV analysis favours s = $\bar{s}$



ATLAS 2302.00336

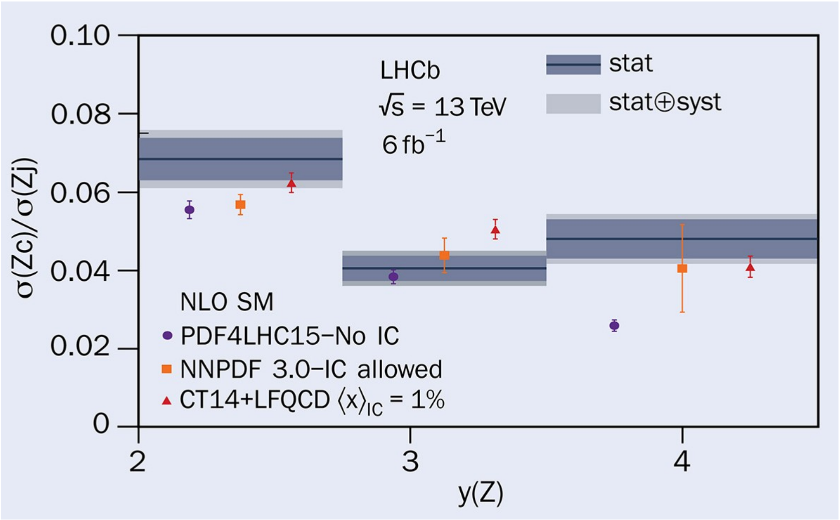
All at NLO QCD
higher order corrections needed to fit properly the PDF

Z + charm jet



Similar to W+charm but for charm PDF

Intrinsic charm component?
Clarification needs
→ higher order corrections
→ charm jet definition



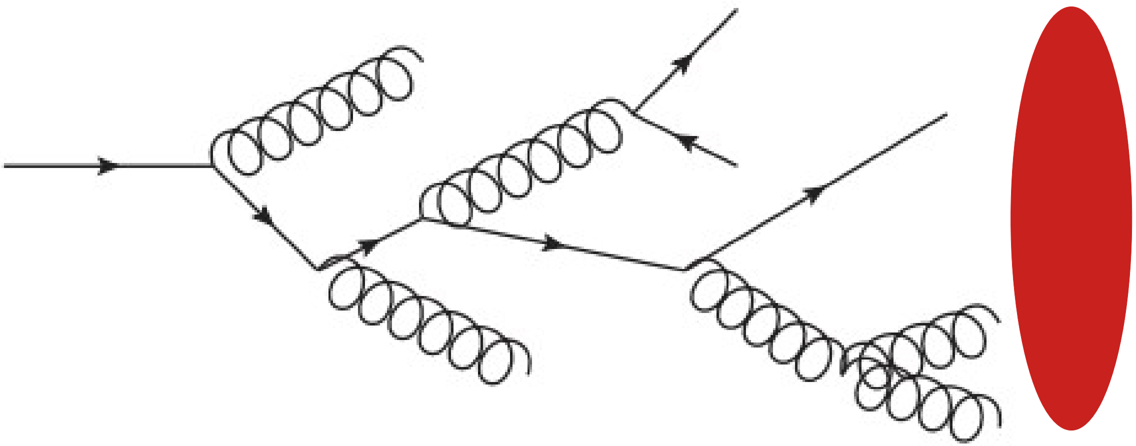
CERN/LHCb 2109.08084

Flavoured jets are everywhere

- Top-quarks
- Vector+heavy flavour: $pp \to W/Z/A + c/b$
- Higgs → charm, Higgs → bottom
- New physics searches
- ...

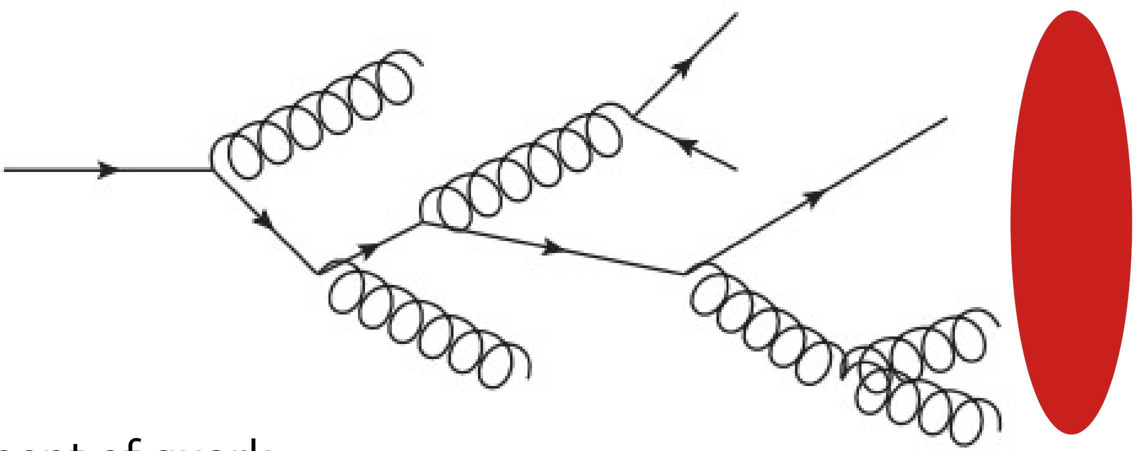
Infrared safety of flavoured jet

Parton evolution



- A high energy parton radiates
- QCD/QED favours collinear splitting
  → jets instead of individual particles
  → closer reconstruction of object from hard interaction
- IR singularities make flavour assignment more difficult

Parton evolution



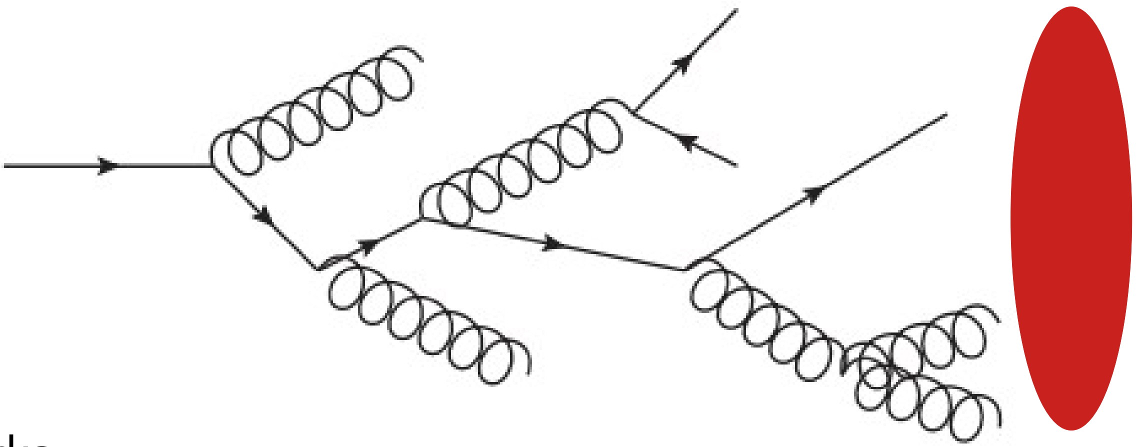
Case I: Massive treatment of quark

- Mass acts as IR regulator → no IR divergences from collinear splitting
- Price to pay: log(pT/m) will be important at high energy!
  → resummation needed for reliable predictions
- Parton-showers can do this but at low accuracy
- Higher order calculations more difficult

} NLO+PS

- Some applications (like PDF fits) need fixed order pQCD at higher orders

Parton evolution



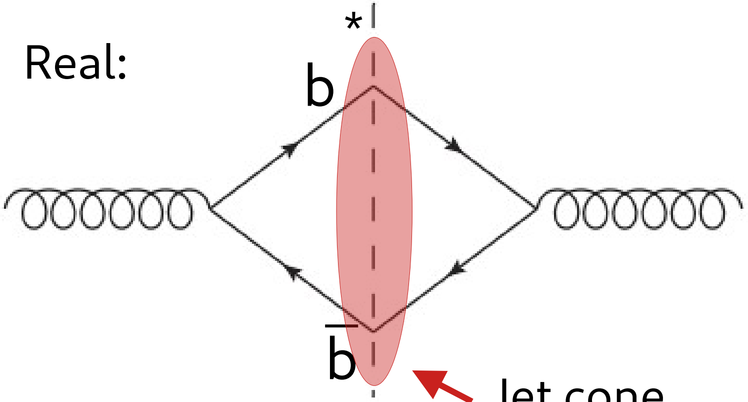
Case II: massless quarks

- Collinear divergences absorbed by renormalisation
- Consistent treatment in junction with PDFs
- Higher order calculations easier → NNLO QCD de-facto standard
- BUT: IR-safety more demanding due to collinear and soft flavoured particles
  → modified algorithms needed
  → implications for phenomenology!

IR safety issues at NLO QCD

Massless QCD: Cancellation of IR divergences between real and virtual corrections

Real:

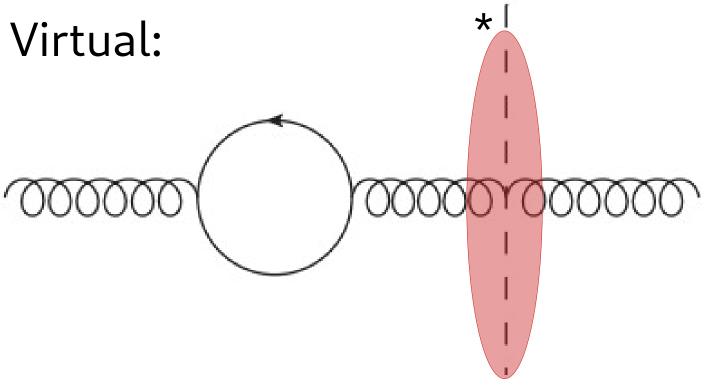


Jet cone

Flavoured?

+

Virtual:



= IR finite

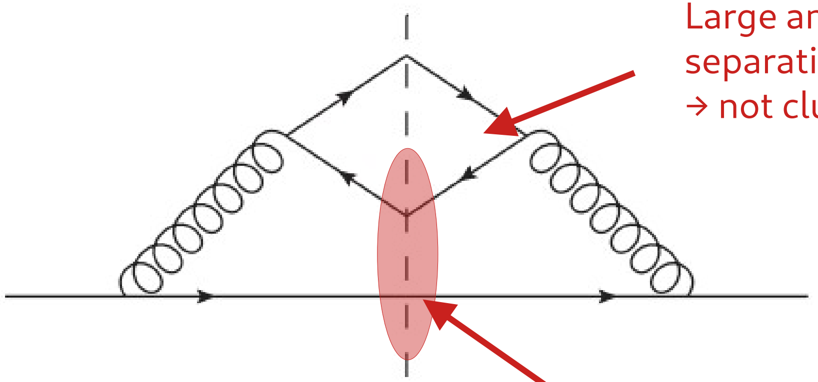
Definitively not flavoured!

$b\,\bar{b}$ has to count as a gluon!

*: cut symbolises the “measured” final state

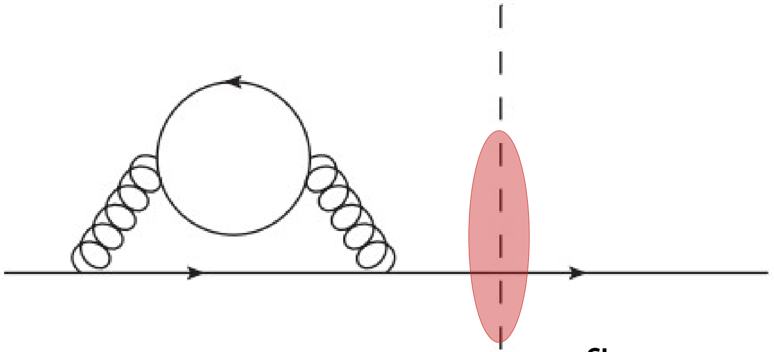
IR safety issues at NNLO QCD

Double soft limit of quark pairs



Large angular separation
→ not clustered

Not flavoured!



flavoured!

- These double soft splitting need to be captured
- Requires to interleave the kinematics and the flavour information!

Solution: Modified jet algorithms

- Implies correlated treatment of kinematics and flavour information

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:

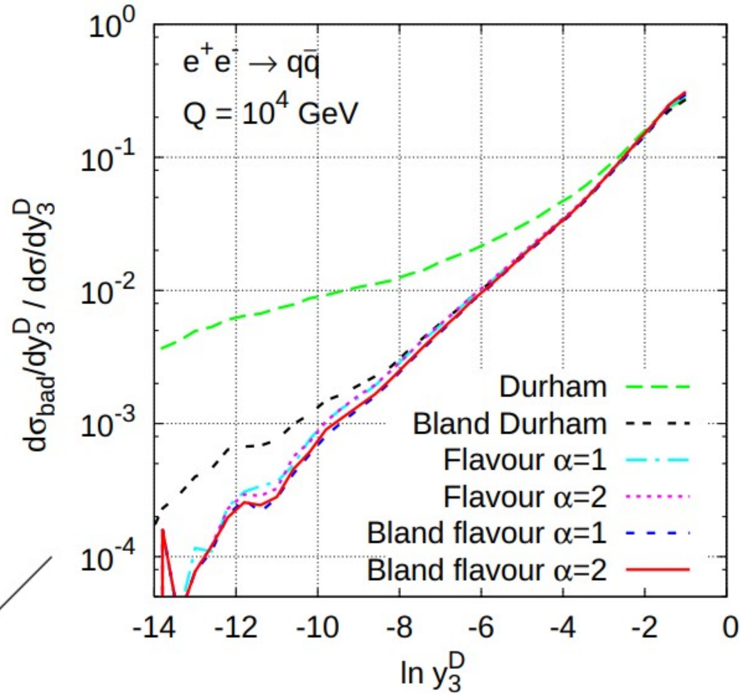
$$d_{i,B} = \begin{cases} \max(k_{T,i}, k_{T,B}(y_i))^\alpha \min(k_{T,i}, k_{T,B}(y_i))^{2-\alpha} & \text{i is flavoured} \\ \min(k_{T,i}, k_{T,B}(y_i))^\alpha & \text{else} \end{cases}$$

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

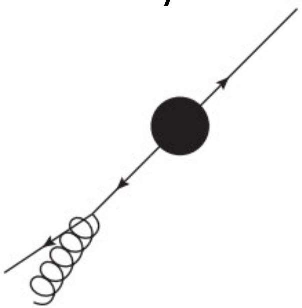
$$d_{\bar{B}}(\eta) = \sum_i k_{T,i}(\theta(\eta - \eta_i) + \theta(\eta_i - \eta)e^{\eta - \eta_i}$$

Tests of IR safety

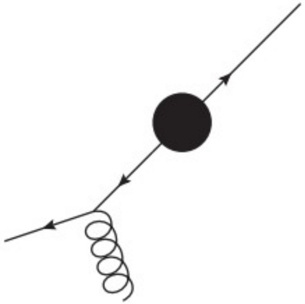
- Rate of bad-identified jet-flavour as a function of IR sensitive variable
- Parton-shower to model many emissions



Small y3



Large y3



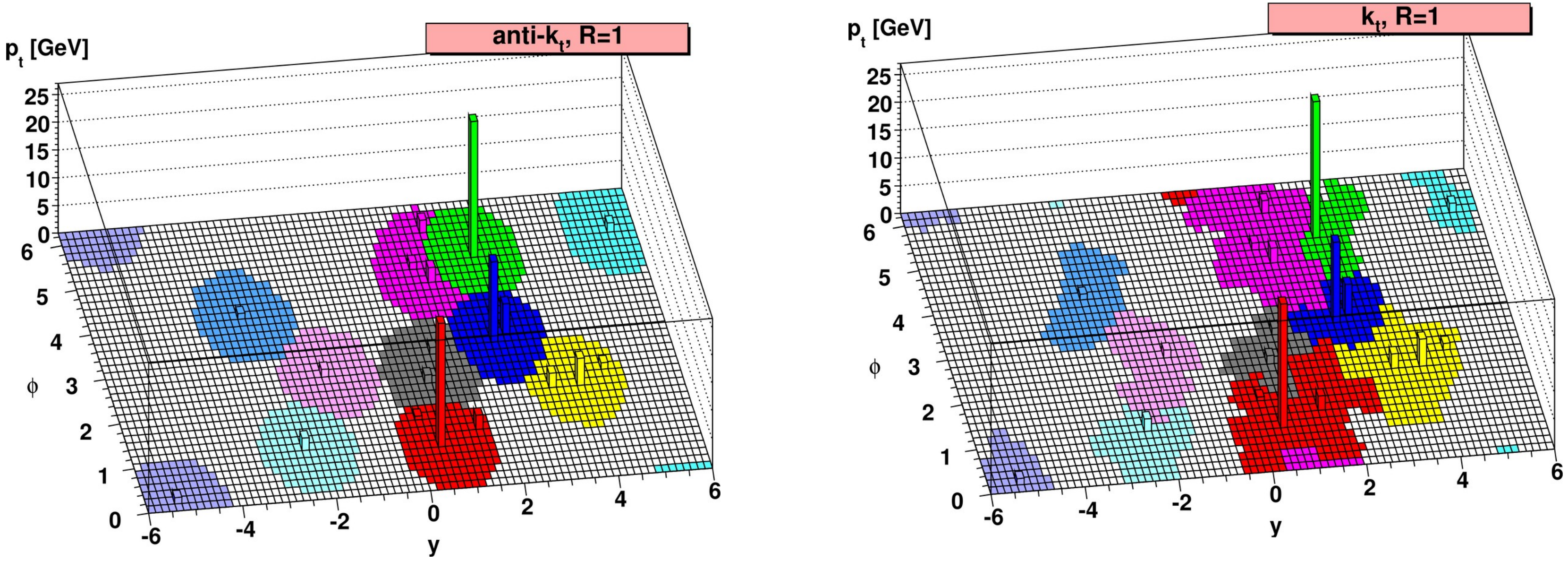
Problem solved, isn't it?

The standard algorithm for the LHC is the anti-kT:

→ nice geometric properties

→ less sensitive to soft physics

Towards Jetography
Salam 0906.1833



New proposals for flavour-safe anti-kT jets

- Flavour anti-kT — **Infrared-safe flavoured anti-kT jets**, Czakon, Mitov, Poncelet 2205.11879
- Flavour with Soft-drop — **Practical Jet Flavour Through NNLO** Caletti, Larkoski, Marzani, Reichelt 2205.01109
- 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
- Flavour neutralisation (not-yet-public)
- TBC…

Flavour anti-kT

Infrared-safe flavoured anti-kT jets, Czakon, Mitov, Poncelet 2205.11879

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

The energy ordering in anti-kT prevents correct recombination of flavoured pairs in the double soft limit.

Proposed modification:

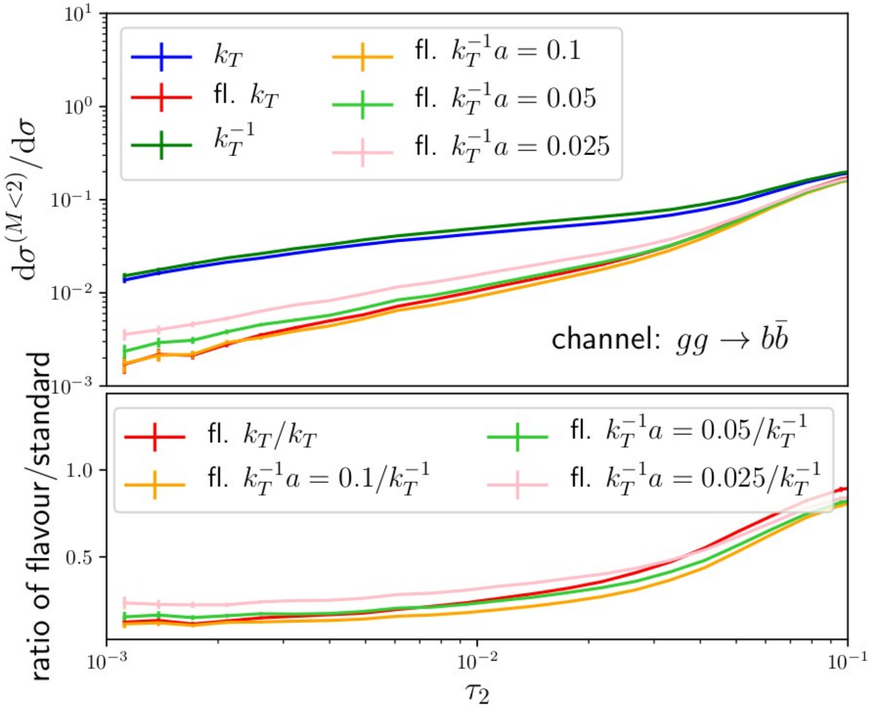
A soft term designed to modify the distance of flavoured pairs.

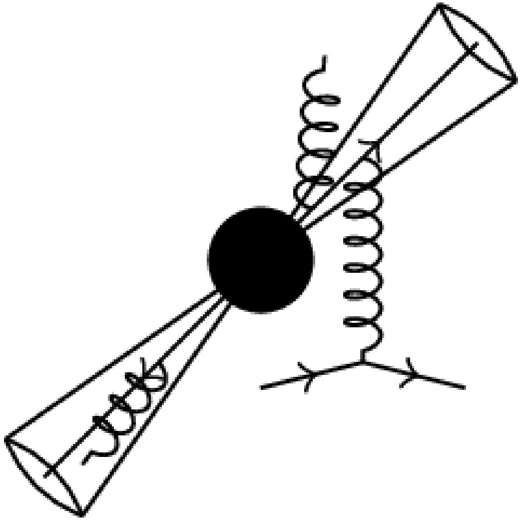
$$d_{ij}^{(F)} = d_{ij} \begin{cases} \mathcal{S}_{ij} & \text{i,j is flavoured pair} \\ 1 & \text{else} \end{cases}$$

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

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



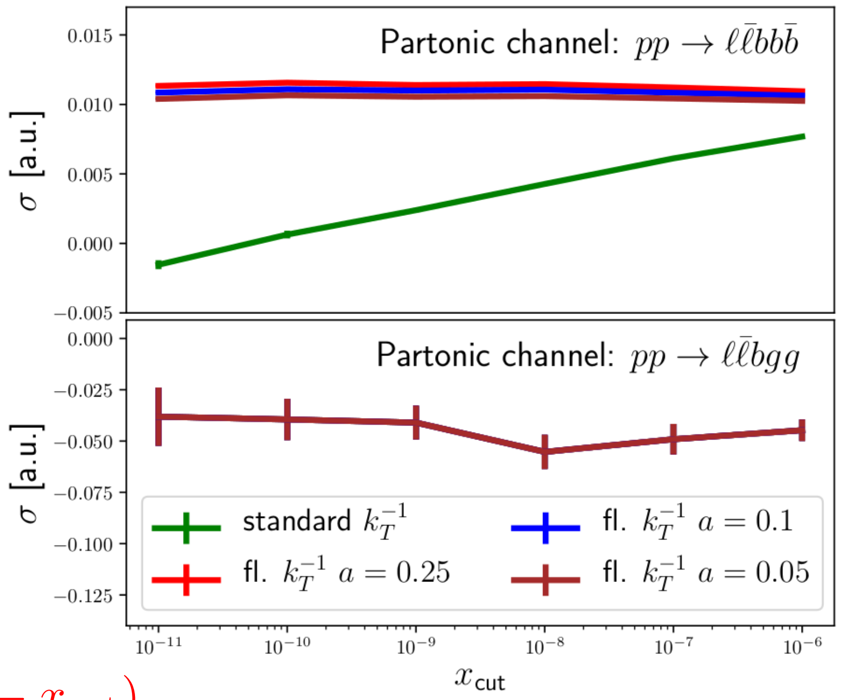


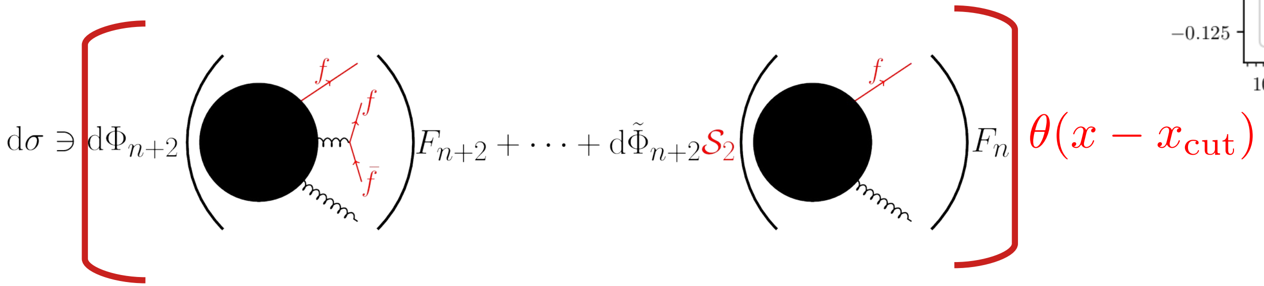
Tests of IR safety with NNLO FO computations

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





Remarks to the flavour anti-kT

$$d_{ij}^{(F)} = d_{ij} \begin{cases} \mathcal{S}_{ij} & \text{i,j is flavoured pair} \\ 1 & \text{else} \end{cases}$$

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

- What is that kT_max parameter?

  Some scale to define what soft means. Examples:
  1. pT of hardest pseudo jet or lepton at a clustering step
  2. Some fixed dynamical scale, e.g. pT(Z), pT(lep), …
  3. Some fixed hard scale: m_top, m_Z etc.

  → The choice impacts the clustering.

Flavour-dressing

A dress of flavour to suit any jet
Gauld, Huss, Stagnitto 2208.11138

- Use standard anti-kT (or any other) to cluster jets {j_k}
- Creation of flavour clusters {f_l}
  → cluster radiation close to quarks (or B/D hadrons)
  → soft-drop criterion for flavour+not-flavour clusters:

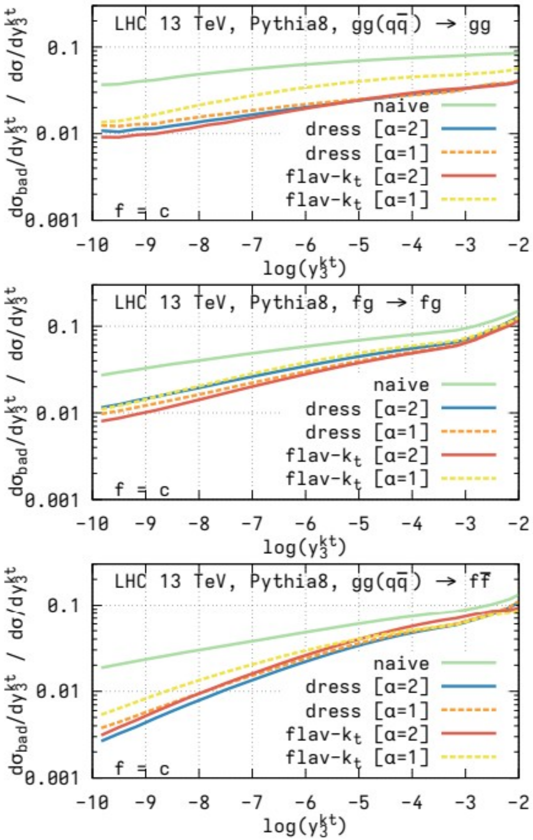
$$\frac{\min(p_{t,a}, p_{t,b})}{(p_{t,a} + p_{t,b})} > z_{\mathrm{cut}} \left(\frac{\Delta R_{ab}}{\delta R}\right)^{\beta}$$

- Associate flavour clusters to jets based on flavour-kT distances

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

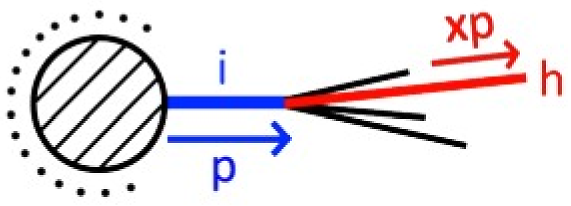
$$d_{i,B} = \begin{cases} \max(k_{T,i}, k_{T,B}(y_i))^{\alpha} \min(k_{T,i}, k_{T,B}(y_i))^{2-\alpha} & \text{i is flavoured} \\ \min(k_{T,i}, k_{T,B}(y_i))^{\alpha} & \text{else} \end{cases}$$

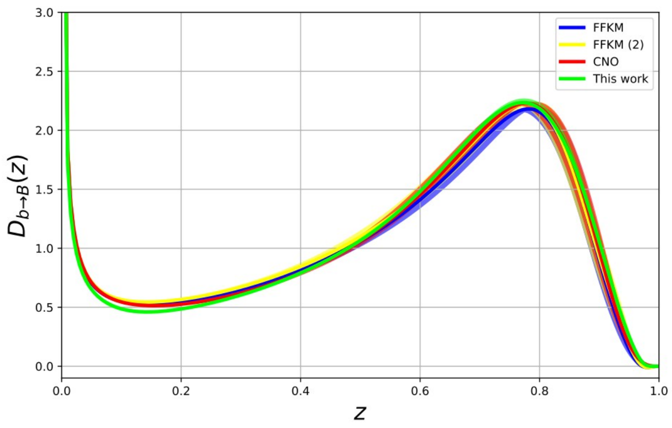
- Accumulation criterion to deal with multiple tags in one jet



Fragmentation

- A massless quark's momentum is ill-defined due to collinear emissions q → q g
- For finite quark momentum (i.e. the quark has "large" momentum) the collinear singularities can be subtracted in the framework of **fragmentation** (similar to PDF)
- Perturbative fragmentation: massless parton to massive quark fragmentation is calculable in pQCD
- Non-perturbative fragmentation into hadrons needs to be fitted to data (like PDFs)





Credit: Terry Generet

Using fragments as tags

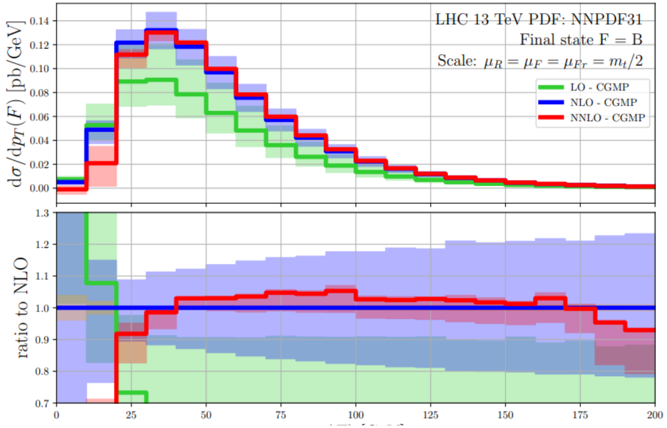
B-hadron production in NNLO QCD: application to LHC ttbar events with leptonic decays, Czakon, Generet, Mitov and Poncelet, 2102.08267

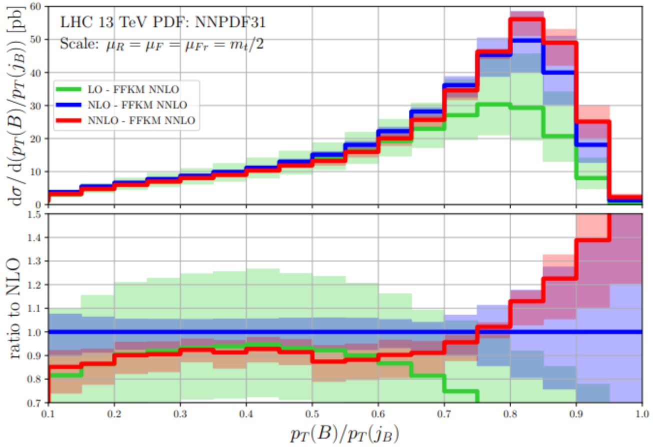
- Advantage is that the fragments momentum is IR safe while the quark's is not → can be used as flavour tag!
- Example:

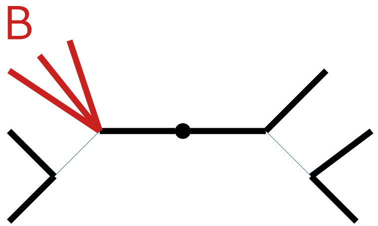
$$pp \to t\bar{t} \to B\ell\bar{\ell}\nu\bar{\nu}b + X$$

typical ttbar selection + B-hadron is part of one jet

- Parton-evolution without parton-shower at high accuracy: Implementation through NNLO QCD







Experimental flavour tagging

Experimental b/c-tagging

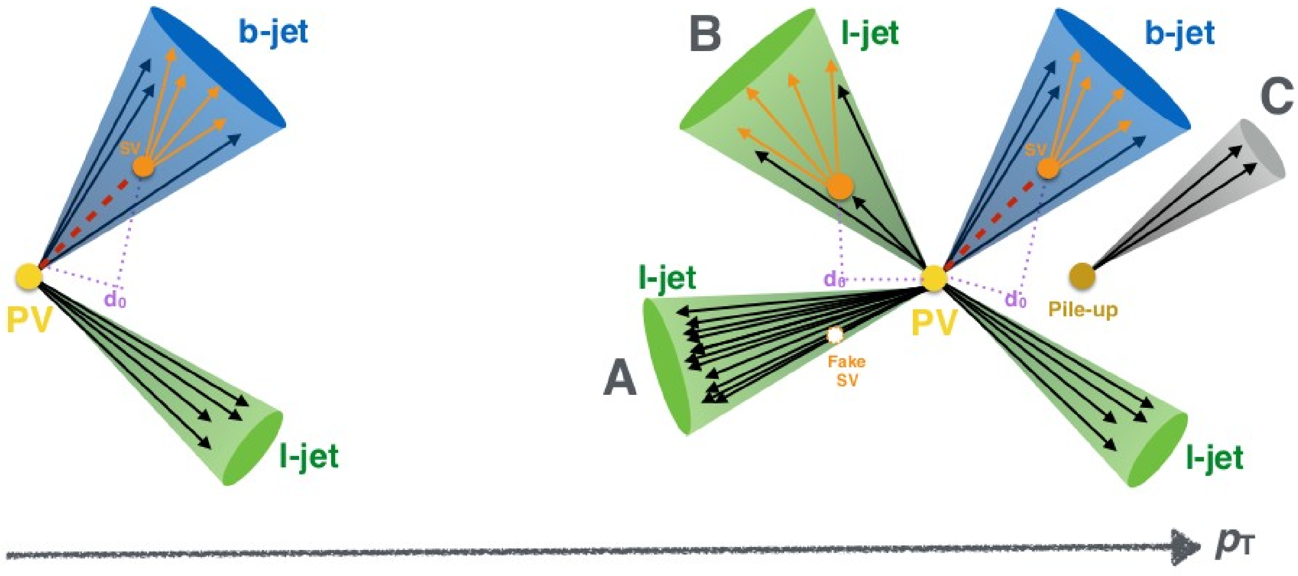
Credit: Arnaud Duperrin (DIS23 talk)

Secondary vertex (SV) tagging

- Long-life time
  → several mm flight
- Looking for the decay products of B-hadron decays forming SV

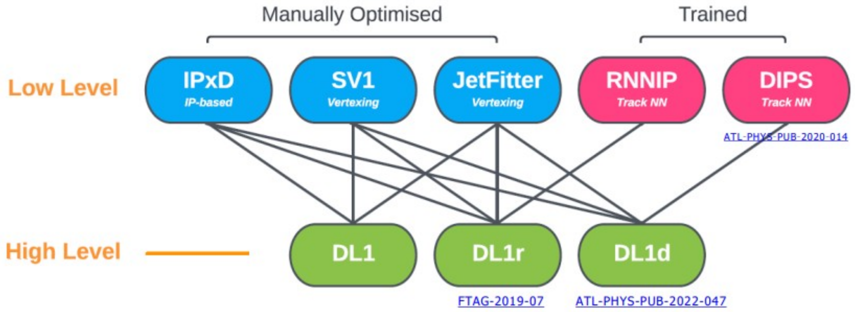
Challenges

- Fake SV from fragmentation
- Material interactions
- Pile-up



Experimental b/c-tagging with NN

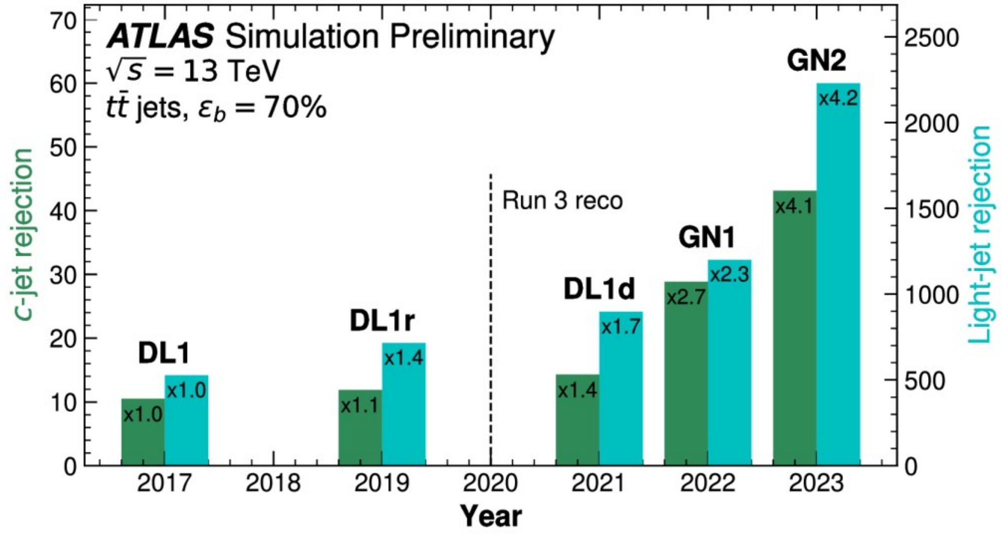
Credit: Arnaud Duperrin (DIS23 talk)



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 from MC simulations



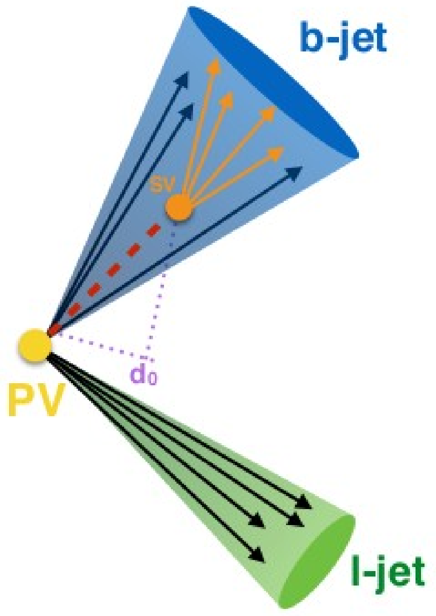
Ghost tagging

A jet is defined as flavoured if:

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

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

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



"Truth" labelling used in Monte Carlo samples, used to train the NN

Technically okay for PS+hadronisation models
BUT
Unsatisfactory from theory point of view

Issues for precision phenomenology

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

- For now: comparisons to higher order QCD partonic computations
  require corrections for the differences in tagging procedures! → Unfolding!

  1) g → b b splitting if both b's hadronise to B-hadrons
  (this is different to b $\bar{b}$ = g @ fixed order)

  2) Hadronisation/non-perturbative models

- Unfolding corrections can be sizeable O(5-10%)

Phenomenology with flavour anti-kT

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)

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

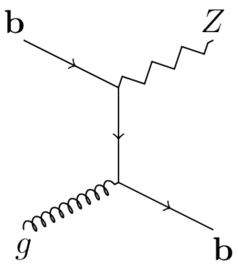
- CMS measurement @ 8 TeV

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

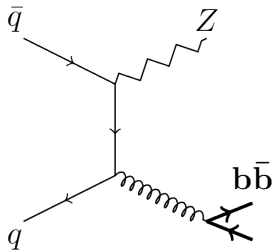
→ Ideal testing ground for flavour anti-kT

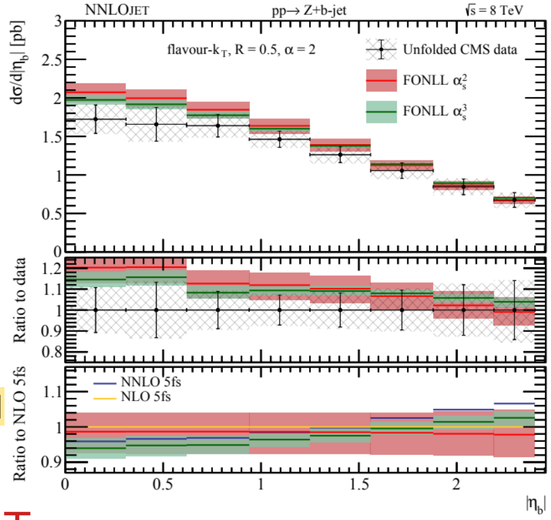
pp → Z(ll) + b-jet

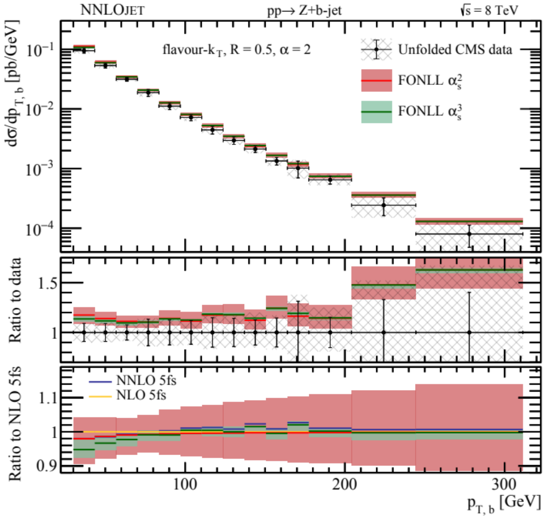
5fs:



4fs:



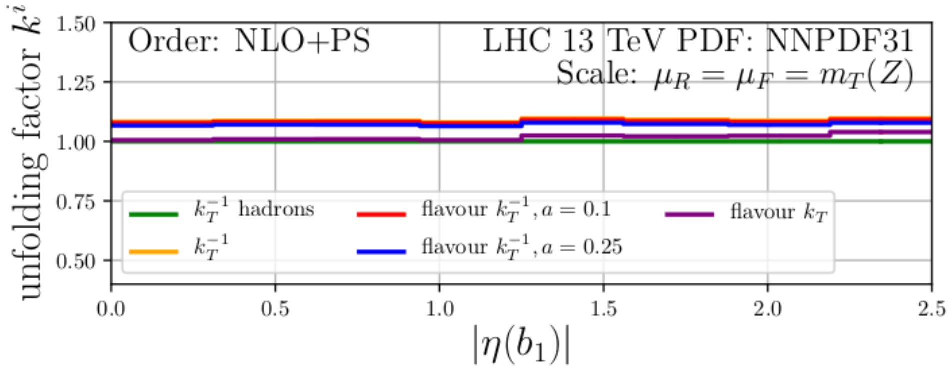


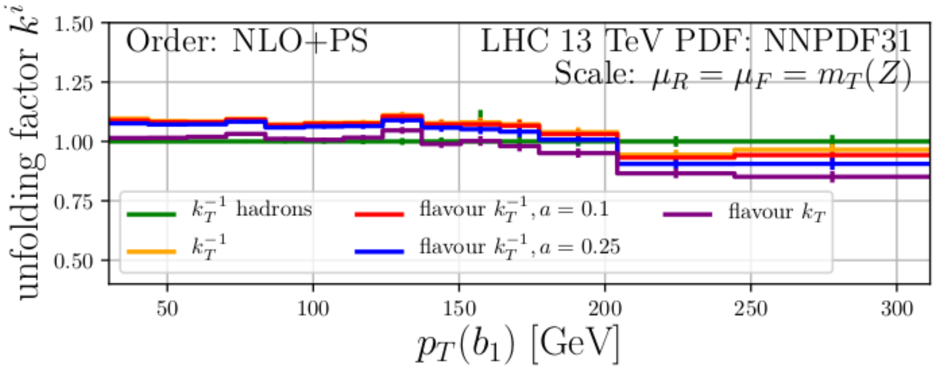


Bin-by-bin unfolding

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

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





Z+b-jet Phenomenology: Tunable parameter

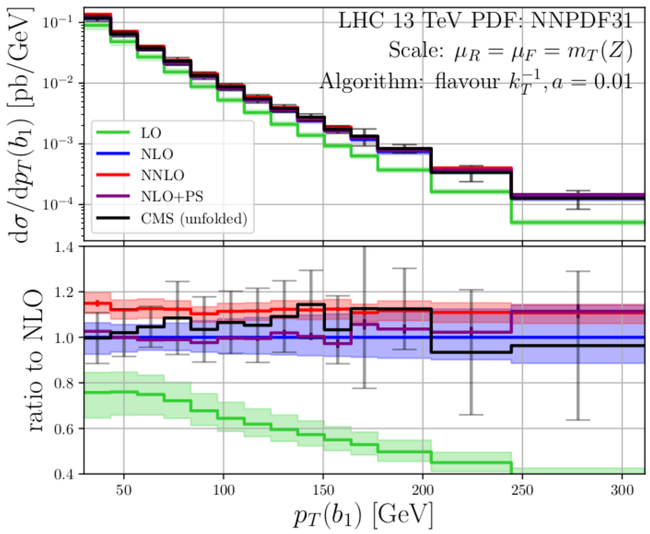
Benchmark process: pp → Z(ll) + b-jet

Tunable parameter a:

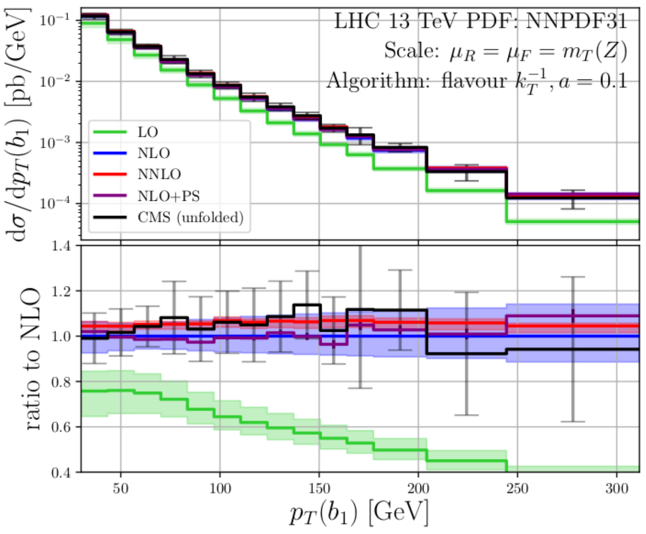
- Limit a → 0 <=> original anti-kT (IR unsafe)
- Large a <=> large modification of cluster sequence

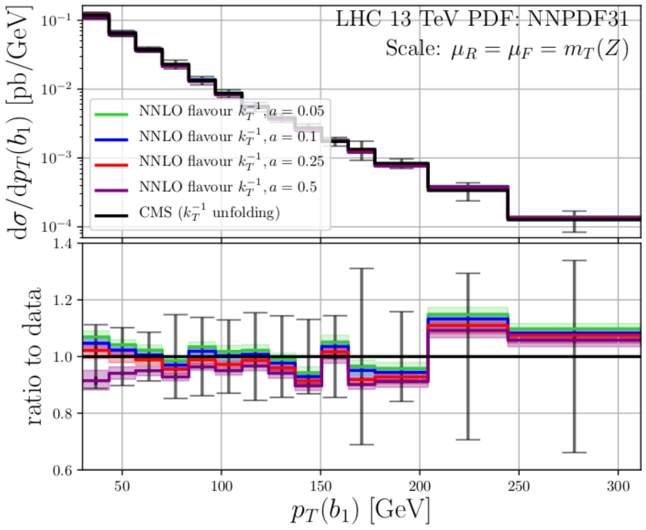
Comparison of different parameter a to data:

Flavour anti-kT (a=0.01):



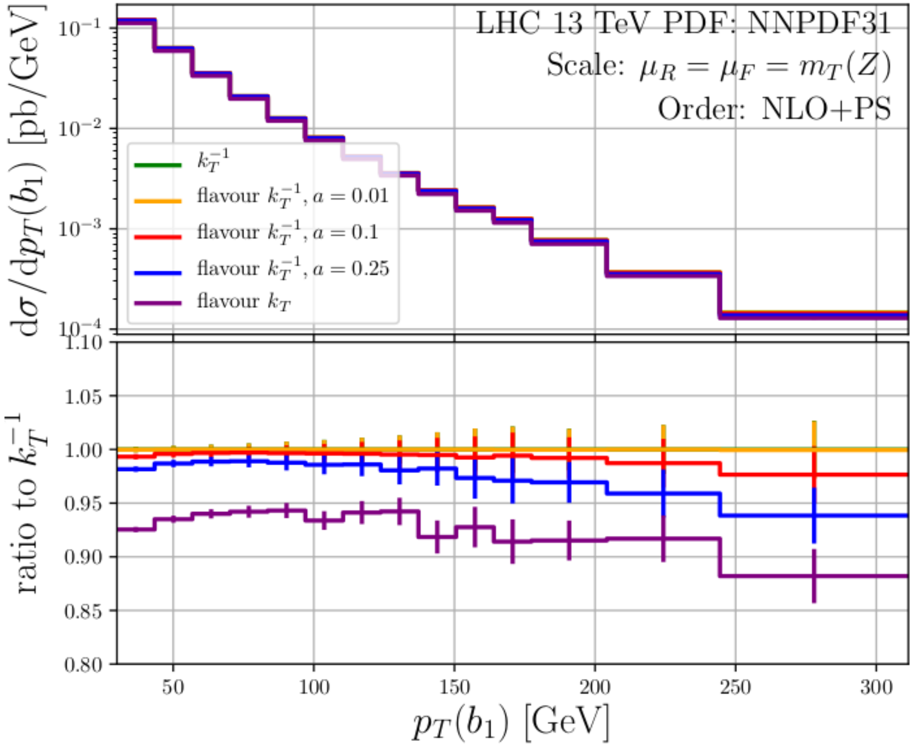
Flavour anti-kT (a=0.1):





Z+b-jet Phenomenology: Tunable parameter II

What happens in the presence of many flavoured partons? → NLO PS



Tunable parameter a:

- Small a: Flavour anti-kT results are more similar to standard anti-kT
- Larger a: Larger modification of clustering

Good FO perturbative convergence + Small difference to standard anti-kT → a~0.1 is a good candidate

Summary

Summary

- Jets are a tool to connect QCD of quarks & gluons to actually strongly interacting particles, i.e. hadrons.
- They are defined by a suitable algorithm: experimentally and theoretically
- “Jet-substructure” reveals additional information:
  - Separation of quark and gluon initiated jets
  - Jets of definite flavour:

Experimentally	*Displayed vertices* of heavy intermediate particles: D/B mesons
MC Event Simulation	Similar objects due to hadronization and detector simulations
Partonic computations	• Impose relation between quarks and hadrons • Massless quarks: emission of soft flavoured pairs → Implications for IR safety in FO computations beyond NLO → Special flavour jet algorithms solve this problem

- Why are partonic computations for flavoured jets interesting?
  - Higher order perturbation theory (not necessarily available matched to PS)
  - Extraction of SM parameters or PDFs

Take home messages

1) Measurements of flavour sensitive observables becoming more precise
   → higher order theory needed to make most of these measurements!

2) Higher order theory predictions require rethinking of jet flavour definitions
   → new algorithms

3) Phenomenology requires joint effort from experiment and theory
   to achieve infra-red safe, precise and accurate theory/data comparisons

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

Top-quark pair final state modelling:

Modeling uncertainties of ttbarW+- multilepton signatures
Bevilacqua, Bi, Cordero, Hartanto , Kraus, Nasufi, Reina, Worek 2109.15181

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