QCD at/for the LHC
Joey Huston
Michigan State University

Di-jet Event at 7 TeV

21 June seminar at Orsay
Some references

Hard interactions of quarks and gluons: a primer for LHC physics

J.M. Campbell, J.W. Huston and W.J. Stirling

1 Department of Physics and Astronomy, University of Glasgow, Glasgow G12 8QQ, UK
2 Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48840, USA
3 Institute for Particle Physics Phenomenology, University of Durham, Durham DH1 3LE, UK

E-mail: j.campbell@physics.gla.ac.uk, huston@msu.edu and w.j.stirling@durham.ac.uk

Received 14 July 2006, in final form 6 November 2006
Published 19 December 2006

Abstract

In this paper, we will develop the perturbative framework for the calculation of hard-scattering processes. We will undertake to provide both a reasonably rigorous development of the formalism of hard-scattering of quarks and gluons as well as an intuitive understanding of the physics behind the scattering. We will emphasize the role of logarithmic corrections as well as power counting in \( \alpha_s \) in order to understand the behaviour of hard-scattering processes. We will include ‘rules of thumb’ as well as ‘official recommendations’, and where possible will seek to dispel some myths. We will also discuss the impact of soft processes on the measurements of hard-scattering processes. Experiences that have been gained at the Fermilab Tevatron will be recounted and, where appropriate, extrapolated to the LHC.

(Some figures in this article are in colour only in the electronic version)

goal is to provide a reasonably global picture of LHC calculations

Review

Jets in hadron–hadron collisions

S.D. Ellis\(^{a,*}\), J. Huston\(^b\), K. Hatakeyama\(^c\), P. Loch\(^d\), M. Tönnesmann\(^e\)

\(^a\) University of Washington, Seattle, WA 98195, United States
\(^b\) Michigan State University, East Lansing, MI 48824, United States
\(^c\) Rockefeller University, New York, NY 10021, United States
\(^d\) University of Arizona, Tucson, AZ 85721, United States
\(^e\) Max Planck Institute für Physik, München, Germany


Abstract

In this article, we review some of the complexities of jet algorithms and of the resultant comparisons of data to theory. We review the extensive experience with jet measurements at the Tevatron, the extrapolation of this acquired wisdom to the LHC and the differences between the Tevatron and LHC environments. We also describe a framework (SpartaJet) for the convenient comparison of results using different jet algorithms.

\(^*\) 2007 Elsevier B.V. All rights reserved.

Keywords: Jet; Jet algorithm; LHC; Tevatron; Perturbative QCD; SpartaJet

Contents

1. Introduction .................................................................................................................. 485
2. Factorization ............................................................................................................... 486
3. Jets: Parton level vs experiment .................................................................................. 490
   3.1. Iterative cone algorithm ...................................................................................... 490
       3.1.1. Definitions .................................................................................................. 490
       3.1.2. \( R_{\text{dep}} \) seeds and IR-sensitivity .............................................................. 495
       3.1.3. Seedless and midpoint algorithms .............................................................. 498
       3.1.4. Merging ..................................................................................................... 499
       3.1.5. Summary .................................................................................................... 499
Towards Jetography

GAVIN P. SALAM

LPTHE, UPMC Univ. Paris 6,
CNRS UMR 7589, 75252 Paris 05, France

Abstract

As the LHC prepares to start taking data, this review is intended to provide a QCD theorist’s understanding and views on jet finding at hadron colliders, including recent developments. My hope is that it will serve both as a primer for the newcomer to jets and as a quick reference for those with some experience of the subject. It is devoted to the questions of how one defines jets, how jets relate to partons, and to the emerging subject of how best to use jets at the LHC.

THE SM AND NLO MULTILEG WORKING GROUP:
Summary Report


1 The University of Edinburgh, School of Physics and Astronomy, Edinburgh EH9 3JZ, UK
2 Institute for Particle Physics, ETH Zurich, CH-8093 Zurich, Switzerland
3 Michigan State University, East Lansing, Michigan 48824, USA
4 Dipartimento di Fisica Teorica e del Cosmo, Centro Autodesk di Fisica di Particelle Elementari (CAFFE), University of Genova, E-16171 Genova, Italy
5 PSI CH-2320 Villigen, Switzerland
6 Dipartimento di Fisica, Polo Scientifico, Università di Torino, I-10125 Torino, Italy
7 Department of Physics, University of California, Berkeley, CA 94720, USA
8 Dipartimento di Fisica, University of Bologna, I-40127 Bologna, Italy
9 Institute for Theoretical Physics, ETH Zurich, CH-8093 Zurich, Switzerland
10 LAM, Université Pierre et Marie Curie, CNRS, F-75252 Paris 05, France
11 Institute of Physics, Polish Academy of Sciences, PL-31342 Cracow, Poland
12 Max Planck Institute for Physics, (Werner-Heisenberg-Institut), D-80893 München, Germany
13 Albert-Ludwigs-Universität Freiburg, Physikalisches Institut, Hermann-Herder-Str. 1, 79104 Freiburg im Breisgau, Germany
14 On leave of absence from INFN, Sez. di Genova, Italy
15 IPP, EFLL, CH-1015 Lausanne, Switzerland
16 University Claude Bernard Lyon I, Institute de Physique Nucleaire de Lyon (IPNL), F-69622 Villeurbanne, CEDEX, France
17 Institute for Theoretical Physics, ETH, CH-8093 Zurich, Switzerland
18 SLAC National Accelerator Laboratory, Stanford University, Stanford, CA 94309, USA
19 Institute of Nuclear Physics, Polish Academy of Sciences, PL-31342 Cracow, Poland
20 Department of Physics, Royal Holloway, University of London, Egham TW20 0EX, UK
Five lectures given for US ATLAS

1. Introduction to pQCD-May 10
2. Higher order calculations-May 17
3. Parton distributions-May 24
4. Jet algorithms-June 1
5. On to the LHC-June 7

Lectures are posted at: www.pa.msu.edu/~huston/atlas/usatlas_lectures
Understanding cross sections at the LHC

- We’re all looking for BSM physics at the LHC
- Before we publish BSM discoveries from the early running of the LHC, we want to make sure that we measure/understand SM cross sections
  - detector and reconstruction algorithms operating properly
  - SM backgrounds to BSM physics correctly taken into account
  - and, in particular, that QCD at the LHC is properly understood
Cross sections at the LHC

- Experience at the Tevatron is very useful, but scattering at the LHC is not necessarily just “rescaled” scattering at the Tevatron.
- Small typical momentum fractions $x$ for the quarks and gluons in many key searches:
  - dominance of gluon and sea quark scattering
  - large phase space for gluon emission and thus for production of extra jets
  - intensive QCD backgrounds
- or to summarize,...lots of Standard Model to wade through to find the BSM pony
Cross sections at the LHC

- Note that the data from HERA and fixed target cover only part of kinematic range accessible at the LHC
- We will access pdf’s down to $10^{-6}$ (crucial for the underlying event) and $Q^2$ up to 100 TeV$^2$
- We can use the DGLAP equations to evolve to the relevant $x$ and $Q^2$ range, but…
  - we’re somewhat blind in extrapolating to lower $x$ values than present in the HERA data, so uncertainty may be larger than currently estimated
  - we’re assuming that DGLAP is all there is; at low $x$ BFKL type of logarithms may become important (more later about DGLAP and BFKL)
Understanding cross sections at the LHC

PDF's, PDF luminosities and PDF uncertainties

Sudakov form factors

underlying event and minimum bias events

jet algorithms and jet reconstruction

LO, NLO and NNLO calculations
K-factors

benchmark cross sections and pdf correlations

I’ll concentrate on PDFs, jets and higher order in this talk.

Most experimenters are/will still mostly use parton shower Monte Carlo for all predictions/theoretical comparisons at the LHC. I’ll try to show that there’s more than that.
Factorization

- Factorization is the key to perturbative QCD
  - the ability to separate the short-distance physics and the long-distance physics
- In the pp collisions at the LHC, the hard scattering cross sections are the result of collisions between a quark or gluon in one proton with a quark or gluon in the other proton
- The remnants of the two protons also undergo collisions, but of a softer nature, described by semi-perturbative or non-perturbative physics

The calculation of hard scattering processes at the LHC requires:
(1) knowledge of the distributions of the quarks and gluons inside the proton, i.e. what fraction of the momentum of the parent proton do they have
-> parton distribution functions (pdf’s)
(2) knowledge of the hard scattering cross sections of the quarks and gluons, at LO, NLO, or NNLO in the strong coupling constant $\alpha_s$
(3) applications of jet algorithms on final state
Hadrons and PDFs

- The proton is a dynamical object; the structure observed depends on the time-scale ($Q^2$) of the observation
- But we know how to calculate this variation (DGLAP)
- We just have to determine the starting points from fits to data

The higher the value of $Q^2$, the more detail we examine

$$f_i(x, Q^2) = \text{number density of partons } i \text{ at momentum fraction } x \text{ and probing scale } Q^2$$
Parton distribution functions and global fits

- Calculation of production cross sections at the LHC relies upon knowledge of pdf’s in the relevant kinematic region.
- Pdf’s are determined by global analyses of data from DIS, DY and jet production.
- Three major groups that provide semi-regular updates to parton distributions when new data/theory becomes available:
  - CTEQ->CTEQ5->CTEQ6 ->CTEQ6.1->CTEQ6.5 ->CTEQ6.6->CT09->CT10
  - NNPDF->NNPDF2.0

Figure 27. The CTEQ6.1 parton distribution functions evaluated at a $Q$ of 10 GeV.
So we have optimal values (minimum $\chi^2$) for the $d=20$ ($22$ for CTEQ6.6, $26$ for CT10) free pdf parameters in the global fit

- $\{a_\mu\}, \mu=1, \ldots d$

Varying any of the free parameters from its optimal value will increase the $\chi^2$

It’s much easier to work in an orthonormal eigenvector space determined by diagonalizing the Hessian matrix, determined in the fitting process

$$H_{\mu\nu} = \frac{1}{2} \frac{\partial \chi^2}{\partial a_\mu \partial a_\nu}$$

To estimate the error on an observable $X(a)$, due to the experimental uncertainties of the data used in the fit, we use the Master Formula

$$\left(\Delta X\right)^2 = \Delta \chi^2 \sum_{\mu,\nu} \frac{\partial X}{\partial a_\mu} \left(H^{-1}\right)_{\mu\nu} \frac{\partial X}{\partial a_\nu}$$

derivative of cross section $X$ along $a_\mu$ eigenvector direction
PDF Errors

- Recap: 20 (22,26) eigenvectors with the eigenvalues having a range of >1E6
- Largest eigenvalues (low number eigenvectors) correspond to best determined directions; smallest eigenvalues (high number eigenvectors) correspond to worst determined directions
- Easiest to use Master Formula in eigenvector basis

\[ \Delta X^+_{\text{max}} = \sqrt{\sum_{i=1}^{N} \max(X_i^+ - X_0, X_i^- - X_0)^2}, \]

\[ \Delta X^-_{\text{max}} = \sqrt{\sum_{i=1}^{N} \max(X_0 - X_i^+, X_0 - X_i^-)^2}. \]

To estimate the error on an observable \( X(a) \), from the experimental errors, we use the Master Formula

\[ (\Delta X)^2 = \Delta \chi^2 \sum_{\mu, \nu} \frac{\partial X}{\partial a_\mu} (H^{-1})_{\mu \nu} \frac{\partial X}{\partial a_\nu} \]

where \( X_i^+ \) and \( X_i^- \) are the values for the observable \( X \) when traversing a distance corresponding to the tolerance \( T(=\sqrt{\Delta \chi^2}) \) along the \( i^{th} \) direction.

Figure 28. A schematic representation of the transformation from the pdf parameter basis to the orthonormal eigenvector basis.
PDF Errors

- What is the tolerance T?
- This is one of the most controversial questions in global pdf fitting?
- We have 2794 data points in the CTEQ6.6 data set (on order of 2000 for CTEQ6.1)
- Technically speaking, a 1-sigma error corresponds to a tolerance T(=sqrt(\Delta\chi^2))=1
- This results in far too small an uncertainty from the global fit
  - with data from a variety of processes from a variety of experiments from a variety of accelerators
- For CTEQ6.1/6.6, we chose a \Delta\chi^2 of 100 to correspond to a 90% CL limit
  - with an appropriate scaling for the larger data set for CTEQ6.6
- In the past, MSTW has chosen a \Delta\chi^2 of 50 for the same limit so CTEQ errors were larger than MSTW errors
  - with new MSTW2008 prescription, errors are closer to CTEQ

\[
\Delta X^+_{\text{max}} = \sqrt{\sum_{i=1}^{N} \left[ \max(X^+_i - X_0, X^-_i - X_0, 0) \right]^2}
\]

\[
\Delta X^-_{\text{max}} = \sqrt{\sum_{i=1}^{N} \left[ \max(X_0 - X^+_i, X_0 - X^-_i, 0) \right]^2}
\]

Figure 29. The pdf errors for the CDF inclusive jet cross section in Run 1 for the 20 different eigenvector directions. The vertical axes show the fractional deviation from the central prediction and the horizontal axes the jet transverse momentum in GeV.
It’s been shown by Jon Pumplin (arXiv: 0909.5176) that a large part of the need for a large value of $\Delta \chi^2$ is because of remaining parameterization biases present even with a very flexible parameterization.

Comparisons with NNPDF (which has less, but not zero, bias) even more important.
Aside: PDF re-weighting

Any physical cross section at a hadron-hadron collider depends on the product of the two pdf's for the partons participating in the collision convoluted with the hard partonic cross section.

Nominally, if one wants to evaluate the pdf uncertainty for a cross section, this convolution should be carried out 41 times (for CTEQ6.1); once for the central pdf and 40 times for the error pdf's.

However, the partonic cross section is not changing, only the product of the pdf's.

So one can evaluate the full cross section for one pdf (the central pdf) and then evaluate the pdf uncertainty for a particular cross section by taking the ratio of the product of the pdf's (the pdf luminosity) for each of the error pdf's compared to the central pdf's.

\[
\sigma_{AB} = \int dx_a dx_b \ f_{a/A}(x_a, Q^2) f_{b/B}(x_b, Q^2) \hat{\sigma}_{ab \rightarrow X}
\]

\[f^i\] is the error pdf and \(f^0\) the central pdf.

\[
\frac{f^i_{a/A}(x_a, Q^2)f^i_{b/B}(x_b, Q^2)}{f^0_{a/A}(x_a, Q^2)f^0_{b/B}(x_b, Q^2)}
\]

This works exactly for fixed order calculations and works well enough for parton shower Monte Carlo calculations.

Most experiments now have code to easily do this…
and many programs will do it for you (MCFM)
A very useful tool

Allows easy calculation and comparison of pdf’s
Uncertainties get large at high $x$.

Uncertainty for gluon larger than that for quarks.

PDF's from one group don't necessarily fall into uncertainty band of another …would be nice if they did.
Uncertainties and parametrizations

- Beware of extrapolations to $x$ values smaller than data available in the fits, especially at low $Q^2$
- Parameterization may artificially reduce the apparent size of the uncertainties
- Compare for example uncertainty for the gluon at low $x$ from the recent neural net global fit to global fits using a parametrization

$Q^2=2 \text{ GeV}^2$

Note: gluon can range negative at low $x$
Correlations

- Consider a cross section $X(a)$
- $i^{th}$ component of gradient of $X$ is
  \[ \frac{\partial X}{\partial a_i} = \partial_i X = \frac{1}{2} (X_i^+ - X_i^-) \]
- Now take 2 cross sections $X$ and $Y$
  - or one or both can be pdf's
- Consider the projection of gradients of $X$ and $Y$ onto a circle of radius 1 in the plane of the gradients in the parton parameter space
- The circle maps onto an ellipse in the $XY$ plane
- The angle $\phi$ between the gradients of $X$ and $Y$ is given by
  \[ \cos \phi = \frac{\nabla X \cdot \nabla Y}{\Delta X \Delta Y} = \frac{1}{4 \Delta X \Delta Y} \sum_{i=1}^{N} \left( X_i^+ - X_i^- \right) \left( Y_i^+ - Y_i^- \right) \]
- The ellipse itself is given by
  \[ \left( \frac{\delta X}{\Delta X} \right)^2 + \left( \frac{\delta Y}{\Delta Y} \right)^2 - 2 \left( \frac{\delta X}{\Delta X} \right) \left( \frac{\delta Y}{\Delta Y} \right) \cos \phi = \sin^2 \phi \]

- If two cross sections/pdf's are very correlated, then $\cos \phi \approx 1$
- ...uncorrelated, then $\cos \phi \approx 0$
- ...anti-correlated, then $\cos \phi \approx -1$

This will prove to be useful
Correlations

- Consider a cross section $X(a)$
- $i^{th}$ component of gradient of $X$ is
  \[
  \frac{\partial X}{\partial a_i} \equiv \partial_i X = \frac{1}{2} (X_i^+ - X_i^-)
  \]
- Now take 2 cross sections $X$ and $Y$
  - or one or both can be pdf's
- Consider the projection of gradients of $X$ and $Y$ onto a circle of radius 1 in the plane of the gradients in the parton parameter space
- The circle maps onto an ellipse in the XY plane
- The angle $\phi$ between the gradients of $X$ and $Y$ is given by
  \[
  \cos \phi = \frac{\vec{\nabla} X \cdot \vec{\nabla} Y}{\Delta X \Delta Y} = \frac{1}{4\Delta X \Delta Y} \sum_{i=1}^{N} \left( X_i^+ - X_i^- \right) \left( Y_i^+ - Y_i^- \right)
  \]
- The ellipse itself is given by
  \[
  \left( \frac{\delta X}{\Delta X} \right)^2 + \left( \frac{\delta Y}{\Delta Y} \right)^2 - 2 \left( \frac{\delta X}{\Delta X} \right) \left( \frac{\delta Y}{\Delta Y} \right) \cos \phi = \sin^2 \phi
  \]

If two cross sections/pdf's are very correlated, then $\cos \phi \approx 1$
- ...uncorrelated, then $\cos \phi \approx 0$
- ...anti-correlated, then $\cos \phi \approx -1$

We'll come back to these later.
Another uncertainty is that due to the variation in the value of $\alpha_s$.

MSTW has recently tried to better quantify the uncertainty due to the variation of $\alpha_s$, by performing global fits over a finer range, taking into account correlations between the values of $\alpha_s$ and the PDF errors.

Procedure is a bit complex.

PDF uncertainties given by [1]

$$(\Delta F_{\text{PDF}})^{\alpha_s}_{+} = \sqrt{\sum_{k=1}^{n} \left( \max \left( F_{\text{PDF}}(S_k^+), F_{\text{PDF}}(S_k^-) \right) - F_{\text{PDF}}(S_0) \right)^2},$$

$$(\Delta F_{\text{PDF}})^{\alpha_s}_{-} = \sqrt{\sum_{k=1}^{n} \left( \max \left( F_{\text{PDF}}(S_0), F_{\text{PDF}}(S_k^+) \right) - F_{\text{PDF}}(S_k^-) \right)^2},$$

for each of the five fixed values of $\alpha_s$. Then the overall best-fit prediction is $F_{\text{PDF}}^{\alpha_s}(S_0)$, and the overall "PDF+$\alpha_s$" uncertainties are given by

$$(\Delta F_{\text{PDF} + \alpha_s})^{+} = \max_{\alpha_s} \left( \{ F^{\alpha_s}(S_0) + (\Delta F_{\text{PDF}})^{\alpha_s}_{+} \} - F^{\alpha_s}(S_0) \right),$$

$$(\Delta F_{\text{PDF} + \alpha_s})^{-} = F^{\alpha_s}(S_0) - \min_{\alpha_s} \left( \{ F^{\alpha_s}(S_0) - (\Delta F_{\text{PDF}})^{\alpha_s}_{-} \} \right).$$

Since this prescription might look quite complicated at first sight, we will give a few concrete examples of its application and consequences in the following subsections.$^3$

In many cases, no simple scaling between 68% and 90% CL.
\( \alpha_s(m_Z) \) and uncertainty: a complication

- Different values of \( \alpha_s \) and of its uncertainty are used
- CTEQ and NNPDF use the world average (actually 0.118 for CTEQ and 0.119 for NNPDF), where MSTW2008 uses 0.120, as determined from their best fit

Latest world average (from Siggi Bethke->PDG)

\[ \alpha_s(m_Z) = 0.1184 \pm 0.0007 \]

- What does the error represent?
  - Siggi said that only one of the results included in his world average was outside this range
  - suppose we say that \( \pm 0.002 \) is a reasonable estimate of the uncertainty
\( \alpha_s(m_Z) \) and uncertainty

- Could it be possible for all global PDF groups to use the world average value of \( \alpha_s \) in their fits, plus a prescribed 90\% range for its uncertainty (if not 0.002, then perhaps another acceptable value)?
- After that, world peace
- For the moment, we try determining uncertainties from \( \alpha_s \) over a range of +/- 0.002 from the central value for each PDF group; we also calculate cross sections with a common value of \( \alpha_s = 0.119 \) for comparison purposes
Cross sections should be calculated with MSTW2008, CTEQ6.6 and NNPDF.

- Upper range of prediction should be given by upper limit of error prediction using prescription for combining $\alpha_s$ uncertainty with error PDFs:
  - in quadrature for CTEQ6.6 and NNPDF.
  - using eigenvector sets for different values of $\alpha_s$ for MSTW2008.
  - note that this effectively creates a larger $\alpha_s$ uncertainty range.

- Ditto for lower limit.

- So for a Higgs mass of 120 GeV at 14 TeV, it turns out that the gg cross section lower limit would be defined by the CTEQ6.6 lower limit (PDF+$\alpha_s$ error) and the upper limit defined by the MSTW2008 upper limit (PDF+$\alpha_s$ error).
  - with the difference between the central values primarily due to $\alpha_s$.
  - I’ll come back to using the Higgs as an example in the last lecture.

- To fully understand similarities/differences of cross sections/uncertainties conduct a benchmarking exercise, to which all groups are invited to participate.
NNLO addendum

- NNLO is important for some cross sections (as we saw for gg->Higgs)
- Not all processes used for global fits are available at NNLO (inclusive jet production for example)
- Only global fit at NNLO currently is MSTW
- Current paradigm is to apply NLO uncertainty band to NNLO predictions from MSTW
  - basically a factor of 2 increase over MSTW errors by themselves

- Most of NNLO corrections for Higgs production are from matrix element rather than differences in PDFs between NLO and NNLO
- So K factor (NNLO/LO) can also be used to some reasonable approximation
For CTEQ: $\alpha_s$ series

- Take CTEQ6.6 as base, and vary $\alpha_s(m_Z) +/-0.002$ (in 0.001 steps) around central value of 0.118
- Blue is the PDF uncertainty from eigenvectors; green is the uncertainty in the gluon from varying $\alpha_s$
- We have found that change in gluon due to $\alpha_s$ error (+/-0.002 range) is typically smaller than PDF uncertainty with a small correlation with PDF uncertainty over this range
  - as shown for gluon distribution on right
- **PDF error and $\alpha_s$ error can be added in quadrature**
  - expected because of small correlation
  - in recent CTEQ paper, it has been proven this is correct regardless of correlation, within quadratic approximation to $\chi^2$ distribution

arXiv:1004.4624; PDFs available from LHAPDF

So the CTEQ prescription for calculating the total uncertainty (PDF+$\alpha_s$) involves the use of the 45 CTEQ6.6 PDFs and the two extreme $\alpha_s$ error PDF’s (0.116 and 0.120)

This also means that one can naively scale between 68% and 90% CL.
PDF Benchmarking Exercise 2010

Benchmark processes, all to be calculated

(i) at NLO (in MSbar scheme)
(ii) in 5-flavour quark schemes (definition of scheme to be specified)
(iii) at 7 TeV [ and 14 TeV] LHC
(iv) for central value predictions and +-68%cl [and +- 90%cl] pdf uncertainties
(v) and with +- $\alpha_s$ uncertainties
(vi) repeat with $\alpha_s(m_Z)=0.119$

(prescription for combining with pdf errors to be specified)

Using (where processes available) MCFM 5.7
- gzipped version prepared by John Campbell using the specified parameters and exact input files for each process (and the new CTEQ6.6 $\alpha_s$ series)->thanks John!
- sent out on first week of March (and still available to any interested parties)
- statistics ok for total cross section comparisons
Cross Sections

1. $W^+$, $W^-$, and $Z$ total cross sections and rapidity distributions total cross section ratios $W^+/W^-$ and $(W^+ + W^-)/Z$, rapidity distributions at $y = -4, -3, ..., +4$ and also the $W$ asymmetry: $A_W(y) = \frac{dW^+}{dy} - \frac{dW^-}{dy}(dW^+/dy + dW^-/dy)$ using the following parameters taken from PDG 2009
   - $M_Z = 91.188$ GeV
   - $M_W = 80.398$ GeV
   - zero width approximation
   - $G_F = 0.116637 \times 10^{-5}$ GeV$^{-2}$
   - other EW couplings derived using tree level relations
   - $\text{BR}(Z \rightarrow l\bar{l}) = 0.03366$
   - $\text{BR}(W \rightarrow l\nu) = 0.1080$
   - CKM mixing parameters from eq.(11.27) of PDG2009 CKM review
     
     \[
     V_{\text{CKM}} = \begin{pmatrix}
     0.97419 & 0.2257 & 0.00359 \\
     0.2256 & 0.97334 & 0.0415 \\
     0.00874 & 0.0407 & 0.999133
     \end{pmatrix}
     \]
   - scales: $\mu_R = \mu_F = M_Z$ or $M_W$
Cross Sections

2. gg->H total cross sections at NLO
   - $M_H = 120, 180$ and $240 \text{ GeV}$
   - zero Higgs width approximation, no BR
   - top loop only, with $m_{\text{top}} = 171.3 \text{ GeV}$ in $\sigma_0$
   - scales: $\mu_R = \mu_F = M_H$

3. ttbar total cross section at NLO
   - $m_{\text{top}} = 171.3 \text{ GeV}$
   - zero top width approximation, no BR
   - scales: $\mu_R = \mu_F = m_{\text{top}}$
Some results from the benchmarking

- ...from G. Watt’s presentation at PDF4LHC meeting on March 26
- See also S. Glazov's summary in the March 31 MC4LHC workshop at CERN
- CTEQ/MSTW predictions for $W$ cross section/uncertainty in very good agreement
  - small impact from different $\alpha_s$ value
  - similar uncertainty bands
- NNPDF prediction low because of use of ZM-VFNS
- HERAPDF1.0 a bit high because of use of combined HERA dataset
W/Z ratio

- Good agreement among the PDF groups
Higgs cross sections and $\alpha_s$

- A linear dependence of Higgs cross section at NLO with $\alpha_s$ can be observed.
- $\alpha_s$ and gluon distribution are anti-correlated in this range ($\alpha_s$ goes up, gluon goes down), but the Higgs cross section has a large K-factor (NLO/LO), so $\alpha_s$ dependence comes from the higher order contribution.

---

Figure 7: Cross section for Higgs production (mass = 120 GeV) at the LHC, with center of mass energy 7 TeV and 14 TeV, as a function of $\alpha_s(M_Z)$. The predictions for the 45 alternative PDF sets are shown for $\alpha_s(M_Z) = 0.118$. For the other values of $\alpha_s(M_Z)$ (= 0.116, 0.117, 0.119, 0.120) only the central prediction is shown. The combined uncertainty range (CTEQ6.6+CT66AS) is shown as the error bar; cf. Table 1.
Some results from the benchmarking

- …from G. Watt’s presentation at PDF4LHC meeting on March 26
- Similar gluon-gluon luminosity uncertainty bands, as noted before
- Cross sections fall into two groups, outside 68% CL error bands
- But, slide everyone’s prediction along the $\alpha_s$ curve to 0.119 (for example) and predictions agree reasonably well
  - within 68% CL PDF errors
More benchmarking
New from CTEQ-TEA (Tung et al)->CT10 PDFs

- Combined HERA-1 data
- CDF and D0 Run-2 inclusive jet data
- Tevatron Run 2 Z rapidity from CDF and D0
- W electron asymmetry from CDFII and D0II (D0 muon asymmetry) (in CT10W)
- Other data sets same as CTEQ6.6
- All data weights set to unity (except for CT10W)
- Tension observed between D0 II electron asymmetry data and NMC/BCDMS data
- Tension between D0 II electron and muon asymmetry data

- Experimental normalizations are treated on same footing as other correlated systematic errors
- More flexible parametrizations: 26 free parameters (26 eigenvector directions)
- Dynamic tolerance: look for 90% CL along each eigenvector direction
  - within the limits of the quadratic approximation, can scale between 68% and 90% CL with naïve scaling factor
- Two series of PDF’s are introduced
  - CT10: no Run 2 W asymmetry
  - CT10W: Run 2 W asymmetry with an extra weight
CT10/CT10W predictions

No big changes with respect to CTEQ6.6

Total cross sections

$\sigma(W^+)/\sigma(W^-)$ vs. $y_W$ at the LHC

$\sigma(W^\pm)/\sigma(Z^0)$ vs. $y_{W/Z}$ at the LHC
Define a correlation cosine between two quantities $Z_t$ and $tT$.

If two cross sections are very correlated, then $\cos \varphi \approx 1$

...uncorrelated, then $\cos \varphi \approx 0$

...anti-correlated, then $\cos \varphi \approx -1$

Figure 1: Dependence on the correlation ellipse formed in the $\Delta X - \Delta Y$ plane on the value of the correlation cosine $\cos \varphi$. 
Correlations with Z, tT

Define a correlation cosine between two quantities.

- If two cross sections are very correlated, then $\cos \phi \approx 1$
- ...uncorrelated, then $\cos \phi \approx 0$
- ...anti-correlated, then $\cos \phi \approx -1$

- Note that correlation curves to Z and to tT are mirror images of each other.

- By knowing the pdf correlations, can reduce the uncertainty for a given cross section in ratio to a benchmark cross section \textit{iff} $\cos \phi > 0$; e.g. $\Delta (\sigma_{W^+} + \sigma_{Z}) / \sigma_Z \approx 1\%$
- If $\cos \phi < 0$, pdf uncertainty for one cross section normalized to a benchmark cross section is larger
- So, for gg$\rightarrow$H(500 GeV); pdf uncertainty is 4%; $\Delta (\sigma_{H} / \sigma_Z) \approx 8\%$
NLO cross sections

- Less sensitivity to unphysical input scales, i.e. renormalization and factorization scales
- First level of prediction where normalization (and sometimes shape) can be taken seriously
- More physics
  - parton merging gives structure in jets
  - initial state radiation
  - more species of incoming partons
- Suppose I have a cross section $\sigma$ calculated to NLO ($O(\alpha_s^n)$)
- Any remaining scale dependence is of one order higher ($O(\alpha_s^{n+1})$)
  - in fact, we know the scale dependent part of the $O(\alpha_s^{n+1})$ cross section before we perform the complete calculation, since the scale-dependent terms are explicit at the previous order

$$\frac{d\sigma}{dE_T} = \alpha_s(\mu_R)^2 A + \alpha_s(\mu_R)^3 (B + 2b_0 L A) + \alpha_s(\mu_R)^4 (C + 3b_0 L B + (3b_0^2 L^2 + 2b_1 L) A)$$

with $L = \log(\mu_R/E_T)$ and $b_i$ the known beta function coefficients.

Inclusive jet prod at NNLO

we know A and B, not C

LO has monotonic scale dependence

non-monotonic at NLO

Figure 11: Single jet inclusive distribution at $E_T = 100$ GeV and $0.1 < |\eta| < 0.7$ at $\sqrt{s} = 1800$

The NNLO coefficient $C$ is unknown. The curves show the guesses $C = 0$ (solid) and $C = \pm B^2/A$ (dashed).
### MCFM

- Many processes available at LO and NLO
  - note these are partonic level only
- Option for ROOT output; PDF errors automatically calculated
- mcfm.fnal.gov

<table>
<thead>
<tr>
<th>Process</th>
<th>Process</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p\bar{p} \rightarrow W^\pm /Z$</td>
<td>$p\bar{p} \rightarrow W^+ + W^-$</td>
</tr>
<tr>
<td>$p\bar{p} \rightarrow W^\pm + Z$</td>
<td>$p\bar{p} \rightarrow Z + Z$</td>
</tr>
<tr>
<td>$p\bar{p} \rightarrow W^\pm + \gamma$</td>
<td>$p\bar{p} \rightarrow W^\pm /Z + H$</td>
</tr>
<tr>
<td>$p\bar{p} \rightarrow W^\pm + g^* \rightarrow b\bar{b}$</td>
<td>$p\bar{p} \rightarrow Z b\bar{b}$</td>
</tr>
<tr>
<td>$p\bar{p} \rightarrow W^\pm /Z + 1 \text{ jet}$</td>
<td>$p\bar{p} \rightarrow W^\pm /Z + 2 \text{ jets}$</td>
</tr>
<tr>
<td>$pp(gg) \rightarrow H$</td>
<td>$pp(gg) \rightarrow H + 1 \text{ jet}$</td>
</tr>
<tr>
<td>$pp(VV) \rightarrow H + 2 \text{ jets}$</td>
<td>$p\bar{p} \rightarrow t + X$</td>
</tr>
<tr>
<td>$pp \rightarrow t + W$</td>
<td>(2 jets now)</td>
</tr>
</tbody>
</table>
**State of the art**

<table>
<thead>
<tr>
<th>Relative order</th>
<th>2-&gt;1</th>
<th>2-&gt;2</th>
<th>2-&gt;3</th>
<th>2-&gt;4</th>
<th>2-5</th>
<th>2-&gt;6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>LO</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_s$</td>
<td>NLO</td>
<td>LO</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_s^2$</td>
<td>NNLO</td>
<td>NLO</td>
<td>LO</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_s^3$</td>
<td></td>
<td>NNLO</td>
<td>NLO</td>
<td>LO</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_s^4$</td>
<td></td>
<td></td>
<td>NLO</td>
<td>LO</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_s^5$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LO</td>
<td></td>
</tr>
</tbody>
</table>

- **LO**: well under control, even for multiparticle final states
- **NLO**: well understood for 2->1, 2->2 and 2->3; first calculations of 2->4 (W +3 jets, ttbb)
- **NNLO**: known for inclusive and exclusive 2->1 (i.e. Higgs, Drell-Yan); work on 2->2 (Higgs + 1 jet)
### An experimenter’s wishlist

<table>
<thead>
<tr>
<th>Single Boson</th>
<th>Diboson</th>
<th>Triboson</th>
<th>Heavy Flavour</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W^+ \leq 5j$</td>
<td>$WW^+ \leq 5j$</td>
<td>$WWW^+ \leq 3j$</td>
<td>$t\bar{t}^+ \leq 3j$</td>
</tr>
<tr>
<td>$W + b\bar{b} \leq 3j$</td>
<td>$W + b\bar{b}^+ \leq 3j$</td>
<td>$WWWW + b\bar{b}^+ \leq 3j$</td>
<td>$t\bar{t} + \gamma^+ \leq 2j$</td>
</tr>
<tr>
<td>$W + c\bar{c} \leq 3j$</td>
<td>$W + c\bar{c}^+ \leq 3j$</td>
<td>$WWW + \gamma\gamma^+ \leq 3j$</td>
<td>$t\bar{t} + W^+ \leq 2j$</td>
</tr>
<tr>
<td>$Z^+ \leq 5j$</td>
<td>$ZZ^+ \leq 5j$</td>
<td>$Z\gamma\gamma^+ \leq 3j$</td>
<td>$t\bar{t} + Z^+ \leq 2j$</td>
</tr>
<tr>
<td>$Z + b\bar{b}^+ \leq 3j$</td>
<td>$Z + b\bar{b}^+ \leq 3j$</td>
<td>$ZZZ^+ \leq 3j$</td>
<td>$t\bar{t} + H^+ \leq 2j$</td>
</tr>
<tr>
<td>$Z + c\bar{c}^+ \leq 3j$</td>
<td>$ZZ + c\bar{c}^+ \leq 3j$</td>
<td>$WZZ^+ \leq 3j$</td>
<td>$t\bar{b} \leq 2j$</td>
</tr>
<tr>
<td>$\gamma^+ \leq 5j$</td>
<td>$\gamma\gamma^+ \leq 5j$</td>
<td>$ZZZ^+ \leq 3j$</td>
<td>$b\bar{b}^+ \leq 3j$</td>
</tr>
<tr>
<td>$\gamma + b\bar{b} \leq 3j$</td>
<td>$\gamma\gamma + b\bar{b} \leq 3j$</td>
<td></td>
<td>single top</td>
</tr>
<tr>
<td>$\gamma + c\bar{c} \leq 3j$</td>
<td>$\gamma\gamma + c\bar{c} \leq 3j$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$WZ^+ \leq 5j$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$WZ + b\bar{b} \leq 3j$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$WZ + c\bar{c} \leq 3j$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$W\gamma^+ \leq 3j$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$Z\gamma^+ \leq 3j$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Realistic wishlist

- Was developed at Les Houches in 2005, and expanded in 2007 and 2009
- Calculations that are important for the LHC AND do-able in finite time
- In 2009, we added tttt, Wbbj, Z+3j, W+4j plus an extra column for each process indicating the level of precision required by the experiments to see for example if EW corrections may need to be calculated
- In order to be most useful, decays for final state particles (t,W,H) need to be provided in the codes as well
- Since the publication of Les Houches 2009 in March, processes 6 and 7 have been completed
- V + 4 jets (process 10) is on the horizon

<table>
<thead>
<tr>
<th>Process ((V \in {Z, W, \gamma}))</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calculations completed since Les Houches 2005</td>
<td></td>
</tr>
<tr>
<td>1. (pp \rightarrow VV) jet</td>
<td>(WW)jet completed by Dittmaier/Kallweit/Uwer [4, 5]: Campbell/Ellis/Zanderighi [6]. (ZZ)jet completed by Binoth/Griesberg/Kang/Kaminski/Sangunni [7]</td>
</tr>
<tr>
<td>2. (pp \rightarrow Higgs+2)jets</td>
<td>NLO QCD to the (gg) channel completed by Campbell/Ellis/Zanderighi [8]; NLO QCD+EW to the VBF channel completed by Ciccolini/Dennen/Dittmaier [9, 10]</td>
</tr>
<tr>
<td>3. (pp \rightarrow VVV)</td>
<td>(ZZZ) completed by Lazopoulos/Melnikov/Petriello [11] and (WWZ) by Hankele/Zeppenfeld [12] (see also Binoth/Ossola/Papadopoulos/Pittau [13])</td>
</tr>
<tr>
<td>4. (pp \rightarrow t\bar{t}) jets</td>
<td>relevant for (t\bar{t}H) computed by Breidenstein/Dennen/Dittmaier/Pozzorini [14, 15] and Bevilacqua/Czakon/Papadopoulos/Pittau/Worek [16] calculated by the Blackhat/Sherpa [17] and RooKit [18] collaborations</td>
</tr>
<tr>
<td>5. (pp \rightarrow V+3)jets</td>
<td></td>
</tr>
<tr>
<td>Calculations remaining from Les Houches 2005</td>
<td></td>
</tr>
<tr>
<td>6. (pp \rightarrow t\bar{t}+2)jets</td>
<td>relevant for (t\bar{t}H) computed by Bevilacqua/Czakon/Papadopoulos/Worek [19]</td>
</tr>
<tr>
<td>7. (pp \rightarrow VV)</td>
<td>relevant for VBF (\rightarrow H \rightarrow VV, t\bar{t}H)</td>
</tr>
<tr>
<td>8. (pp \rightarrow VV+2)jets</td>
<td>relevant for VBF (\rightarrow H \rightarrow VV)</td>
</tr>
<tr>
<td>VBF contributions calculated by (Bozzi/Siger/Oleari/Zeppenfeld) [20–22]</td>
<td></td>
</tr>
<tr>
<td>NLO calculations added to list in 2007</td>
<td></td>
</tr>
<tr>
<td>9. (pp \rightarrow b\bar{b})</td>
<td>(q\bar{q}) channel calculated by Golern collaboration [23]</td>
</tr>
<tr>
<td>NLO calculations added to list in 2009</td>
<td></td>
</tr>
<tr>
<td>10. (pp \rightarrow V+4) jets</td>
<td>top pair production, various new physics signatures</td>
</tr>
<tr>
<td>11. (pp \rightarrow W+b)</td>
<td>top, new physics signatures</td>
</tr>
<tr>
<td>12. (pp \rightarrow t\bar{t})</td>
<td>various new physics signatures</td>
</tr>
<tr>
<td>Calculations beyond NLO added in 2007</td>
<td></td>
</tr>
<tr>
<td>13. (gg \rightarrow W+V+O(\alpha_s^2))</td>
<td>backgrounds to Higgs</td>
</tr>
<tr>
<td>14. NNLO (pp \rightarrow t\bar{t})</td>
<td>normalization of a benchmark process</td>
</tr>
<tr>
<td>15. NNLO to VBF and (Z/\gamma)+jet</td>
<td>Higgs couplings and SM benchmark</td>
</tr>
<tr>
<td>Calculations including electroweak effects</td>
<td></td>
</tr>
<tr>
<td>16. NNLO QCD+NNLO EW for (W/Z)</td>
<td>precision calculation of a SM benchmark</td>
</tr>
</tbody>
</table>

Table 1: The updated experimenter’s wishlist for LHC processes
Loops and legs

2->4 is very impressive

but just compare to the complexity of the sentences that Sarah Palin uses.
Some issues/questions

- Once we have the calculations, how do we (experimentalists) use them?
- Best is to have NLO partonic level calculation interfaced to parton shower/hadronization
  - but that has been done only for relatively simple processes and is very (theorist) labor intensive
    - still waiting for inclusive jets in MC@NLO, for example

- Even with partonic level calculations, need public code and/or ability to write out ROOT ntuples of parton level events
  - so that can generate once with loose cuts and distributions can be re-made without the need for the lengthy re-running of the predictions
  - what is done for example with MCFM for CTEQ4LHC
  - it’s what Blackhat+Sherpa has provided me for \( W + 3 \) jets at NLO
    - but 10’s of Gbytes for file sizes
Proposed common ntuple output

- A generalization of the FROOT format used in MCFM
- Writeup in NLM proceedings

<table>
<thead>
<tr>
<th>ROOT Tree Branch</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Npart/I</td>
<td>number of partons (incoming and outgoing)</td>
</tr>
<tr>
<td>Px[Npart]/D</td>
<td>Px of partons</td>
</tr>
<tr>
<td>Py[Npart]/D</td>
<td>Py of partons</td>
</tr>
<tr>
<td>Pz[Npart]/D</td>
<td>Pz of partons</td>
</tr>
<tr>
<td>E[Npart]/D</td>
<td>E of partons</td>
</tr>
<tr>
<td>x1/D</td>
<td>Bjorken-x of incoming parton 1</td>
</tr>
<tr>
<td>x2/D</td>
<td>Bjorken-x of incoming parton 2</td>
</tr>
<tr>
<td>id1/I</td>
<td>PDG particle ID of incoming parton 1</td>
</tr>
<tr>
<td>id2/I</td>
<td>PDG particle ID of incoming parton 2</td>
</tr>
<tr>
<td>fac_scale/D</td>
<td>factorization scale</td>
</tr>
<tr>
<td>ren_scale/D</td>
<td>renormalization scale</td>
</tr>
<tr>
<td>weight/D</td>
<td>global event weight</td>
</tr>
<tr>
<td>Nuwgt/I</td>
<td>number of user weights</td>
</tr>
<tr>
<td>user_wqts[Nuwgt]/D</td>
<td>user event weights</td>
</tr>
<tr>
<td>evt_nc/L</td>
<td>unique event number (identifier)</td>
</tr>
<tr>
<td>Nptr/I</td>
<td>number of event pointers</td>
</tr>
<tr>
<td>evt_pointers[Nptr]/L</td>
<td>event pointers (identifiers of related events)</td>
</tr>
<tr>
<td>Npdfs/I</td>
<td>number of PDF weights</td>
</tr>
<tr>
<td>pdf_wqts[Npdfs]/D</td>
<td>PDF weights</td>
</tr>
</tbody>
</table>

LhaNLOEvent* evt = new LhaNLOEvent();
evtx->addParticle(px1,py1,pz1,E1);
evtx->setProcInfo(x1,id1,x2,id2);
evtx->setRenScale(scale);
...

Another class LhaNLOTreeIO is responsible for writing the events into the ROOT tree and outputting the tree to disk. In addition to the event-wise information global data such as comments, cross sections etc can be written as well. An example is shown below:

LhaNLOTreeIO* writer = new LhaNLOTreeIO(); // create tree writer
writer->initWrite("’test.root’");
...
writer->writeComment("’4+4 jets at NNLO’"); // write global comments
writer->writeComment("’total cross section: Yts/-IJK fb’");
...
writer->writeEvent(*evt); // write event to tree (in event loop)
...
writer->writeTree(); // write tree to disk

Similarly, a tree can be read back from disk:

LhaNLOTreeIO* reader = new LhaNLOTreeIO(); // init reader
ierr=reader->initRead("test.root");
if (ierr) {
  for (int i=0; i<reader->getNumberOfEvents();i++) {
    event->reset();
    ierr=reader->readEvent(i,*event);
    ...
  }
}
K-factors

- Often we work at LO by necessity (parton shower Monte Carlos), but would like to know the impact of NLO corrections.
- K-factors (NLO/LO) can be a useful short-hand for this information.
- But caveat emptor; the value of the K-factor depends on a number of things:
  - PDFs used at LO and NLO
  - scale(s) at which the cross sections are evaluated
- And often the NLO corrections result in a shape change, so that one K-factor is not sufficient to modify the LO cross sections.
### K-factor table from CHS paper

#### K-factors for LHC
- Slightly less K-factors at Tevatron

#### K-factors with NLO PDFs at LO
- More often closer to unity

#### Shapes of distributions may be different at NLO than at LO, but sometimes it is still useful to define a K-factor.

#### Note the value of the K-factor depends critically on its definition.

<table>
<thead>
<tr>
<th>Process</th>
<th>Typical scales</th>
<th>Tevatron $K$-factor</th>
<th>LHC $K$-factor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$K(\mu_0)$</td>
<td>$K(\mu_1)$</td>
</tr>
<tr>
<td>$W$</td>
<td>$m_W$</td>
<td>2$m_W$</td>
<td>1.33</td>
</tr>
<tr>
<td>$W+1\text{jet}$</td>
<td>$m_W$</td>
<td>$p_T^{\text{jet}}$</td>
<td>1.42</td>
</tr>
<tr>
<td>$W+2\text{jets}$</td>
<td>$m_W$</td>
<td>$p_T^{\text{jet}}$</td>
<td>1.16</td>
</tr>
<tr>
<td>$WW+\text{jet}$</td>
<td>$m_W$</td>
<td>2$m_W$</td>
<td>1.19</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>$m_t$</td>
<td>2$m_t$</td>
<td>1.08</td>
</tr>
<tr>
<td>$t\bar{t}+1\text{jet}$</td>
<td>$m_t$</td>
<td>2$m_t$</td>
<td>1.13</td>
</tr>
<tr>
<td>$b\bar{b}$</td>
<td>$m_b$</td>
<td>2$m_b$</td>
<td>1.20</td>
</tr>
<tr>
<td>Higgs</td>
<td>$m_H$</td>
<td>$p_T^{\text{jet}}$</td>
<td>2.33</td>
</tr>
<tr>
<td>Higgs via VBF</td>
<td>$m_H$</td>
<td>$p_T^{\text{jet}}$</td>
<td>1.07</td>
</tr>
<tr>
<td>Higgs+1jett</td>
<td>$m_H$</td>
<td>$p_T^{\text{jet}}$</td>
<td>2.02</td>
</tr>
<tr>
<td>Higgs+2jett</td>
<td>$m_H$</td>
<td>$p_T^{\text{jet}}$</td>
<td>–</td>
</tr>
</tbody>
</table>

Table 3: $K$-factors for various processes at the LHC calculated using a selection of input parameters. Have to fix this table. In all cases, the CTEQ6M PDF set is used at NLO. $K$ uses the CTEQ6L1 set at leading order, whilst $K'$ uses the same set, CTEQ6M, as at NLO and $K''$ uses the modified LO (2-loop) PDF set. For Higgs+1,2jets, a jet cut of 40 GeV/c and $|\eta| < 4.5$ has been applied. A cut of $p_T^{\text{jet}} > 20$ GeV/c has been applied for the $t\bar{t}+\text{jet}$ process, and a cut of $p_T^{\text{jet}} > 50$ GeV/c for $WW+\text{jet}$. In the $W$ (Higgs)+2jets process the jets are separated by $\Delta R > 0.52$, whilst the VBF calculations are performed for a Higgs boson of mass 120 GeV. In each case the value of the $K$-factor is compared at two often-used scale choices, where the scale indicated is used for both renormalization and factorization scales.
Some rules-of-thumb

NLO corrections are larger for processes in which there is a great deal of color annihilation

- **gg->Higgs**
- **gg->γγ**
- **K(gg->tT) > K(qQ -> tT)**
- these gg initial states want to radiate like crazy (see Sudakovs)

NLO corrections decrease as more final-state legs are added

- **K(gg->Higgs + 2 jets)**
  - **< K(gg->Higgs + 1 jet)**
  - **< K(gg->Higgs)**
- unless can access new initial state gluon channel

Can we generalize for uncalculated HO processes?

What about effect of jet vetoes on K-factors? Signal processes compared to background. Of current interest.

---

<table>
<thead>
<tr>
<th>Process</th>
<th>Typical scales</th>
<th>Tevatron K-factor</th>
<th>LHC K-factor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>μ0</td>
<td>μ1</td>
<td>K(μ0)</td>
</tr>
<tr>
<td>W</td>
<td>mW</td>
<td>2mW</td>
<td>1.33</td>
</tr>
<tr>
<td>W+1jet</td>
<td>mW</td>
<td>pT,W</td>
<td>1.42</td>
</tr>
<tr>
<td>W+2jets</td>
<td>mW</td>
<td>pT,W</td>
<td>1.16</td>
</tr>
<tr>
<td>WW+jet</td>
<td>mW</td>
<td>2mW</td>
<td>1.19</td>
</tr>
<tr>
<td>t¯t</td>
<td>m_t</td>
<td>2m_t</td>
<td>1.08</td>
</tr>
<tr>
<td>tt+1jet</td>
<td>m_t</td>
<td>2m_t</td>
<td>1.13</td>
</tr>
<tr>
<td>bb</td>
<td>m_b</td>
<td>2m_b</td>
<td>1.20</td>
</tr>
<tr>
<td>Higgs</td>
<td>m_H</td>
<td>pT,H</td>
<td>2.33</td>
</tr>
<tr>
<td>Higgs via VBF</td>
<td>m_H</td>
<td>pT,H</td>
<td>1.07</td>
</tr>
<tr>
<td>Higgs+1jet</td>
<td>m_H</td>
<td>pT,H</td>
<td>2.02</td>
</tr>
<tr>
<td>Higgs+2jets</td>
<td>m_H</td>
<td>pT,H</td>
<td>–</td>
</tr>
</tbody>
</table>

Table 2: K-factors for various processes at the Tevatron and the LHC calculated using a selection of input parameters. In all cases, the CTEQ6M PDF set is used at NLO. K uses the CTEQ6L1 set at leading order, whilst K’ uses the same set, CTEQ6M, as at NLO. For most of the processes listed, jets satisfy the requirements pT > 15 GeV/c and |η| < 2.5 (5.0) at the Tevatron (LHC). For Higgs+1,2jets, a jet cut of 40 GeV/c and |η| < 4.5 has been applied. A cut of pT,H > 20 GeV/c has been applied for the t+jet process, and a cut of pT,H > 50 GeV/c for WW+jet. In the W(Higgs)+2jets process the jets are separated by ΔR > 0.5, whilst the VBF calculations are performed for a Higgs boson of mass 120 GeV. In each case the value of the K-factor is compared at two often-used scale choices, where the scale indicated is used for both renormalization and factorization scales.

Casimir for biggest color representation final state can be in

Simplistic rule

\[ C_{i1} + C_{i2} - C_{f,max} \]

L. Dixon

Casimir color factors for initial state
Shape dependence of a K-factor

- Inclusive jet production probes very wide $x, Q^2$ range along with varying mixture of $gg, gq, \text{and} qq$ subprocesses
- PDF uncertainties are significant at high $p_T$
- Over limited range of $p_T$ and $y$, can approximate effect of NLO corrections by K-factor but not in general
  - in particular note that for forward rapidities, K-factor $<<1$
  - LO predictions will be large overestimates
  - this is true for both the Tevatron and for the LHC
Consider the W + 3 jets process

<table>
<thead>
<tr>
<th>Process ($V \in {Z, W, \gamma}$)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calculations completed since Les Houches 2005</td>
<td></td>
</tr>
<tr>
<td>1. $pp \rightarrow VV$jet</td>
<td>$WW$jet completed by Dittrich/Kallweit/Uwer [4, 5]; Campbell/Ellis/Zanderighi [6]. $ZZ$jet completed by Binon/Gleisberg/Kang/Kauer/Sangrainetti [7]</td>
</tr>
<tr>
<td>2. $pp \rightarrow Higgs+2$jets</td>
<td>NLO QCD to the $gg$ channel completed by Campbell/Ellis/Zanderighi [8]; NLO QCD+EW to the VBF channel completed by Ciccioli/Demner/Dittmaier [9, 10]</td>
</tr>
<tr>
<td>3. $pp \rightarrow ZVV$</td>
<td>$ZZZ$ completed by Lazopoulos/Mehlhorn/Petrelli [11] and $WWW$ by Hankel/Zeppenfeld [12] (see also Binosi/Orsooa/Papaefstathiou/Pittau [13])</td>
</tr>
<tr>
<td>4. $pp \rightarrow t\bar{t}b\bar{b}$</td>
<td>relevant for $t\bar{t}H$ computed by Bredenstein/Demner/Dittmaier/Pozzorini [14, 15] and Bevilacqua/Czakon/Papaefstathiou/Pittau/Worek [16] calculated by the BlackHole/Herwig [17] and Rocket [18] collaborations</td>
</tr>
<tr>
<td>5. $pp \rightarrow V+3$jets</td>
<td></td>
</tr>
<tr>
<td>Calculations remaining from Les Houches 2005</td>
<td></td>
</tr>
<tr>
<td>6. $pp \rightarrow t\bar{t}+2$jets</td>
<td>relevant for $t\bar{t}H$ computed by Bevilacqua/Czakon/Papaefstathiou/Worek [19]</td>
</tr>
<tr>
<td>7. $pp \rightarrow VVb\bar{b}$</td>
<td>relevant for $VBF \rightarrow H \rightarrow VV$, $t\bar{t}H$</td>
</tr>
<tr>
<td>8. $pp \rightarrow VV+2$jets</td>
<td>relevant for $VBF \rightarrow H \rightarrow VV$; VBF contributions calculated by (Bozzi)Jager/Oleari/Zeppenfeld [20–22]</td>
</tr>
<tr>
<td>NLO calculations added to list in 2007</td>
<td></td>
</tr>
<tr>
<td>9. $pp \rightarrow b\bar{b}b\bar{b}$</td>
<td>$q\bar{q}$ channel calculated by Golem collaboration [23]</td>
</tr>
<tr>
<td>NLO calculations added to list in 2009</td>
<td></td>
</tr>
<tr>
<td>10. $pp \rightarrow V+4$ jets</td>
<td>top pair production, various new physics signatures</td>
</tr>
<tr>
<td>11. $pp \rightarrow Wb\bar{b}j$</td>
<td>top, new physics signatures</td>
</tr>
<tr>
<td>12. $pp \rightarrow t\bar{t}f$</td>
<td>various new physics signatures</td>
</tr>
<tr>
<td>Calculations beyond NLO added in 2007</td>
<td></td>
</tr>
<tr>
<td>13. $gg \rightarrow W^+W^- O(\alpha_s^2)$</td>
<td>backgrounds to Higgs</td>
</tr>
<tr>
<td>14. NNLO $pp \rightarrow t\bar{t}$</td>
<td>normalization of a benchmark process</td>
</tr>
<tr>
<td>15. NNLO to VBF and $Z/\gamma+1$jet</td>
<td>Higgs couplings and SM benchmark</td>
</tr>
<tr>
<td>Calculations including electroweak effects</td>
<td></td>
</tr>
<tr>
<td>16. NNLO QCD+NLO EW for $W/Z$</td>
<td>precision calculation of a SM benchmark</td>
</tr>
</tbody>
</table>

Table 1: The updated experimenters' wishlist for LHC processes
Now consider $W + 3$ jets

Consider a scale of $m_W$ for $W + 1, 2, 3$ jets. We see the K-factors for $W + 1, 2$ jets in the table below, and recently the NLO corrections for $W + 3$ jets have been calculated, allowing us to estimate the K-factors for that process. It is much smaller than one.

<table>
<thead>
<tr>
<th>Process</th>
<th>Typical scales</th>
<th>Tevatron K-factor</th>
<th>LHC K-factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W$</td>
<td>$m_W$</td>
<td>1.33 1.31 1.21</td>
<td>1.15 1.05 1.15</td>
</tr>
<tr>
<td>$W + 1$jet</td>
<td>$m_W$</td>
<td>1.42 1.20 1.43</td>
<td>1.21 1.32 1.42</td>
</tr>
<tr>
<td>$W + 2$jets</td>
<td>$m_W$</td>
<td>1.36 0.91 1.29</td>
<td>0.89 0.88 1.10</td>
</tr>
<tr>
<td>$W + W$jet</td>
<td>$m_W$</td>
<td>1.19 1.37 1.26</td>
<td>1.33 1.40 1.42</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>$m_t$</td>
<td>1.08 1.31 1.24</td>
<td>1.40 1.59 1.19</td>
</tr>
<tr>
<td>$t\bar{t} + 1$jet</td>
<td>$m_t$</td>
<td>1.13 1.43 1.37</td>
<td>0.97 1.29 1.10</td>
</tr>
<tr>
<td>$bb$</td>
<td>$m_b$</td>
<td>1.20 1.21 2.10</td>
<td>0.98 0.84 2.51</td>
</tr>
<tr>
<td>Higgs</td>
<td>$m_H$</td>
<td>2.33</td>
<td>2.32 1.43</td>
</tr>
<tr>
<td>Higgs via VBF</td>
<td>$p_T^H$</td>
<td>0.97 1.07 1.23</td>
<td>1.34 0.85 0.78</td>
</tr>
<tr>
<td>Higgs+1jet</td>
<td>$p_T^H$</td>
<td>2.02</td>
<td>1.47 1.80 1.33</td>
</tr>
<tr>
<td>Higgs+2jets</td>
<td>$p_T^H$</td>
<td>-</td>
<td>1.90 1.15 1.13</td>
</tr>
</tbody>
</table>

Table 3: K-factors for various processes at the LHC calculated using a selection of input parameters. Have to fix this table. In all cases, the CTEQ6L1 PDF set is used at NLO. K uses the CTEQ6L1 set at leading order, whilst $K'$ uses the same set, CTEQ6M, as at NLO and $K''$ uses the modified LO (2-loop) PDF set. For Higgs+1,2jets, a jet cut of 40 GeV/c and $|\eta| < 4.5$ has been applied. A cut of $p_T^{jet} > 20$ GeV/c has been applied for the $t\bar{t}$+jet process, and a cut of $p_T^{W} > 50$ GeV/c for $WW$+jet. In the $W$(Higgs)+2jets process the jets are separated by $\Delta R > 0.52$, whilst the VBF calculations are performed for a Higgs boson of mass 120 GeV. In each case the value of the K-factor is compared at two often-used scale choices, where the scale indicated is used for both renormalization and factorization scales.

Is the K-factor (at $m_W$) at the LHC surprising?
Is the K-factor (at $m_W$) at the LHC surprising?

The K-factors for $W +$ jets ($p_T > 30$ GeV/c) fall near a straight line, as do the K-factors for the Tevatron. By definition, the K-factors for Higgs + jets fall on a straight line.

Nothing special about $m_W$; just a typical choice.

The only way to know a cross section to NLO, say for $W + 4$ jets or Higgs + 3 jets, is to calculate it, but in lieu of the calculations, especially for observables that we have deemed important at Les Houches, can we understand the behavior with the associated number of jets?

Related to this is:
- understanding the reduced scale dependences/pdf uncertainties for cross section ratios we have been discussing
- scale choices at LO for cross sections uncalculated at NLO
Is the K-factor (at $m_W$) at the LHC surprising?

The K-factors for $W + \text{jets}$ ($p_T > 30 \text{ GeV/c}$) fall near a straight line, as do the K-factors for the Tevatron. By definition, the K-factors for Higgs + jets fall on a straight line.

Nothing special about $m_W$; just a typical choice.

The only way to know a cross section to NLO, say for $W + 4 \text{ jets}$ or Higgs + 3 jets, is to calculate it, but in lieu of the calculations, especially for observables that we have deemed important at Les Houches, can we make rules of thumb?

Related to this is:
- understanding the reduced scale dependences/pdf uncertainties for the cross section ratios we have been discussing
- scale choices at LO for cross sections calculated at NLO
- scale choices at LO for cross sections uncalculated at NLO

To understand this further, we have to discuss jet algorithms.
Jet algorithms at LO

- At (fixed) LO, 1 parton = 1 jet
  - why not more than 1? I have to put a $\Delta R$ cut on the separation between two partons; otherwise, there’s a collinear divergence. LO parton shower programs effectively put in such a cutoff
  - Remember the collinear singularity

$$\log\left(\frac{1}{\Delta R_{34}}\right)$$

- But at NLO, I have to deal with more than 1 parton in a jet, and so now I have to talk about how to cluster those partons
  - i.e. jet algorithms
Jet algorithms at NLO

- At NLO, there can be two partons in a jet, life becomes more interesting and we have to start talking about jet algorithms to define jets
  - the addition of the real and virtual terms at NLO cancels the divergences in each

\[
\begin{align*}
  d_{ij} &= \min\left(p_{T,i}^{2p}, p_{T,j}^{2p}\right) \frac{\Delta R_{ij}^2}{D^2} \\
  d_{ii} &= p_{T,i}^{2p}
\end{align*}
\]

- A jet algorithm is based on some measure of localization of the expected collinear spray of particles
- Start with an inclusive list of particles/partons/calorimeter towers/topoclusters
- End with lists of same for each jet
- ...and a list of particles… not in any jet; for example, remnants of the initial hadrons
- Two broad classes of jet algorithms
  - cluster according to proximity in space: cone algorithms
  - ATLAS uses SISCone
  - cluster according to proximity in momenta: \( k_T \) algorithms
  - ATLAS uses \( k_T, \text{antik}_T \)

Pierre-Antoine Delsart's reverse \( k_T \)
Jet algorithms at LO/NLO

- Remember at LO, 1 parton = 1 jet
- By choosing a jet algorithm with size parameter D, we are requiring any two partons to be > D apart
- The matrix elements have $1/\Delta R$ poles, so larger D means smaller cross sections
  - it’s because of the poles that we have to make a $\Delta R$ cut
- At NLO, there can be two (or more) partons in a jet and jets for the first time can have some structure
  - we don’t need a $\Delta R$ cut, since the virtual corrections cancel the collinear singularity from the gluon emission
  - but there are residual logs that can become important if D is too small
- Also, increasing the size parameter D increases the phase space for including an extra gluon in the jet, and thus increases the cross section at NLO (in most cases)

![Diagram showing jet algorithm parameters](image)

For $D=R_{cone}$, Region I = $k_T$ jets, Region II (nominally) = cone jets; I say nominally because in data not all of Region II is included for cone jets.
Is the K-factor (at $m_W$) at the LHC surprising?

The problem is not the NLO cross section; that is well-behaved. The problem is that the LO cross section sits ‘too-high’. The reason (one of them) for this is that we are ‘too-close’ to the collinear pole ($R=0.4$) leading to an enhancement of the LO cross section (double-enhancement if the gluon is soft (~20 GeV/c)). Note that at LO, the cross section increases with decreasing $R$; at NLO it decreases. The collinear dependence gets stronger as $n_{\text{jet}}$ increases.

The K-factors for $W + 3$ jets would be more normal (>1) if a larger cone size and/or a larger jet $p_T$ cutoff were used. But that’s a LO problem; the best approach is to use the appropriate jet sizes/jet $p_T$’s for the analysis and understand the best scales to use at LO (matrix element + parton shower) to approximate the NLO calculation (as well as comparing directly to the NLO calculation).

For 3 jets, the LO collinear singularity effects are even more pronounced.

NB: here I have used CTEQ6.6 for both LO and NLO; CTEQ6L1 would shift LO curves up.
Don’t believe (fixed) LO predictions for jet cross sections

- Let’s look at predictions for $W + 3$ jets for two different jet algorithms as a function of jet size at the LHC (7 TeV)
- At LO, both antikT and SISCone show a marked decrease in cross section as the jet size increases because of the $\log(1/\Delta R)$ effect
- But at NLO, the two cross sections show little dependence on the jet size, and are similar to each other due to addition of extra gluon in jet possible at NLO
- You’ll see the same thing in ATLAS Monte Carlo

![Graph showing W + 3 jets cross section](image)

Blackhat + Sherpa

Note NLO~LO because a scale of $H_T$ has been used
Scale choices at the Tevatron: $W + \text{jets}$

- At the Tevatron, $m_w$ is a reasonable scale (in terms of K-factor~1)
A scale choice of $m_W$ would be in a region where LO $>>$ NLO. In addition, such a scale choice (or related scale choice), leads to sizeable shape differences in the kinematic distributions. The Blackhat+Sherpa people found that a scale choice of $H_T$ worked best to get a constant K-factor for all distributions that they looked at. Note that from the point-of-view of only NLO, all cross sections with scales above $\sim 100$ GeV seem reasonably stable.
Scale choice: why is $E_T^W$ a bad one at the LHC?

If configuration (a) dominated, then as jet $E_T$ increased, $E_T^W$ would increase along with it. But configuration (b) is kinematically favored for high jet $E_T$'s (smaller partonic center-of-mass energy); $E_T^W$ remains small, and that scale does not describe the process very well.

Note that now split/merge can become important as the partonic jets can overlap and share partons.

Configuration b also tends to dominate in the tails of multi-jet distributions (such as $H_T$ or $M_{ij}$); for high jet $E_T$, $W$ behaves like a massless boson, and so there's a kinematic enhancement when it's soft.

FIG. 9: The $E_T$ distribution of the second jet at LO and NLO, for two dynamical scale choices, $\mu = E_T^W$ (left plot) and $\mu = H_T$ (right plot). The histograms and bands have the same meaning as in previous figures. The NLO distribution for $\mu = E_T^W$ turns negative beyond $E_T = 475$ GeV.
Applying a CKKW-like scale also leads to better agreement for shapes of kinematic distributions. Currently under investigation.

See review of $W + 3$ jets in Les Houches 2009 NLM proceedings.
Choosing jet size

● Experimentally
  ♦ in complex final states, such as $W + n$ jets, it is useful to have jet sizes smaller so as to be able to resolve the $n$ jet structure
  ♦ this can also reduce the impact of pileup/underlying event

● Theoretically
  ♦ hadronization effects become larger as $R$ decreases
  ♦ for small $R$, the $\ln R$ perturbative terms referred to previously can become noticeable
  ♦ this restriction in the gluon phase space can affect the scale dependence, i.e. the scale uncertainty for an $n$-jet final state can depend on the jet size,
  ♦ …under investigation

Another motivation for the use of multiple jet algorithms/parameters (i.e. SpartyJet) in LHC analyses.
Jet sizes and scale uncertainties: the Goldilocks theorem

- Take inclusive jet production at the LHC for transverse momenta of the order of 50 GeV
- Look at the theory uncertainty due to scale dependence as a function of jet size
- It appears to be a minimum for cone sizes of the order of 0.7
  - i.e. if you use a cone size of 0.4, there are residual un-cancelled virtual effects
  - if you use a cone size of 1.0, you are adding too much tree level information with its intrinsically larger scale uncertainty
- This effect becomes smaller for jet $p_T$ values on the order of 100 GeV/c
  - how does it translate for multi-parton final states?
  - currently under investigation
Jets at NLO: more complications

- Construct what is called a Snowmass potential shown in Figure 50, where the towers unclustered into any jet are shaded black. A simple way of understanding these dark towers begins by defining a “Snowmass potential” in terms of the 2-dimensional vector \( \mathbf{r} = (y, \phi) \) via

\[
V(\mathbf{r}) = -\frac{1}{2} \sum_{j} p_{ij} \left( R_{\text{cone}}^2 - (\mathbf{r}_j - \mathbf{r})^2 \right) \Theta \left( R_{\text{cone}}^2 - (\mathbf{r}_j - \mathbf{r})^2 \right).
\]

The flow is then driven by the “force” \( \mathbf{F}(\mathbf{r}) = -\nabla V(\mathbf{r}) \) which is thus given by,

\[
\mathbf{F}(\mathbf{r}) = \sum_{j} p_{ij} \left( \mathbf{r}_j - \mathbf{r} \right) \Theta \left( R_{\text{cone}}^2 - (\mathbf{r}_j - \mathbf{r})^2 \right) = \left( \mathbf{r} C(\mathbf{r}) - \mathbf{r}_j \right) \sum_{j \in C(\mathbf{r})} p_{ij},
\]

where \( \mathbf{r} C(\mathbf{r}) = \left( \mathbf{Y}_C, \mathbf{Z}_C, \right) \) and the sum runs over \( j \in C(\mathbf{r}) \) such that \( \sqrt{(y_j - y)^2 + (\phi_j - \phi)^2} \leq R_{\text{cone}} \). As desired, this force pushes the cone to the stable cone position.

- The minima of the potential function indicates the positions of the stable cone solutions
  - the derivative of the potential function is the force that shows the direction of flow of the iterated cone

- The midpoint solution contains both partons

\[
\frac{(y_j - y)^2 + (\phi_j - \phi)^2}{R_{\text{cone}}^2 - (\mathbf{r}_j - \mathbf{r})^2} \leq \frac{1}{2}.
\]
Jets in real life

- Jets don't consist of 1 fermi partons but have a spatial distribution.
- Can approximate jet shape as a Gaussian smearing of the spatial distribution of the parton energy.
  - The effective sigma ranges between around 0.1 and 0.3 depending on the parton type (quark or gluon) and on the parton $p_T$.
- Note that because of the effects of smearing that:
  - The midpoint solution is (almost always) lost.
    - Thus region II is effectively truncated to the area shown on the right.
  - The solution corresponding to the lower energy parton can also be lost.
    - Resulting in dark towers.

Figure 52. A schematic depiction of the effects of smearing on the midpoint cone jet clustering algorithm.

Figure 22. The parameter space ($dE$) for which two partons will be merged into a single jet.

Figure 50. An example of a Monte Carlo inclusive jet event where the midpoint algorithm has left substantial energy unclustered.

Remember the Snowmass potentials.
In NLO theory, can mimic the impact of the truncation of Region II by including a parameter called $R_{\text{sep}}$

- only merge two partons if they are within $R_{\text{sep}} \times R_{\text{cone}}$ of each other
  - $R_{\text{sep}} \sim 1.3$
- $\sim 4\text{-}5\%$ effect on the theory cross section; effect is smaller with the use of $p_T$ rather than $E_T$
- really upsets the theorists (but there are also disadvantages)

- Dark tower effect is also on order of few (<5)\% effect on the (experimental) cross section
- Dark towers affect every cone algorithm
Don’t believe (fixed) LO predictions for jet cross sections

Compare to ATLAS ALPGEN+ PYTHIA samples (ΔR=0.7 matching so we can only compare to last jet size)

At parton level, antikT is ~25% higher than SISCone (same as we observe here at LO)

At topocluster level, antikT is ~2% higher than SISCone (not the 7% observed here)

Why 2%, not 7%?
Some of the W + 3 parton events reconstructed as 2 jets at the parton level for SISCone are reconstructed as 3 jets at the hadron. The cross section for 3 jets increases.
Try this out in ATLAS Monte Carlo

- Take W + 2 parton events (ALPGEN+PYTHIA), run SISCones 0.7 algorithm on parton level, hadron level (not shown) and topocluster level
- Plot the probability for the two sub-jets to merge as a function of the separation of the original two partons in ΔR
- Color code:
  - red: high probability for merging
  - blue: low probability for merging
- Parton level reconstruction agrees with naïve expectation
- Topocluster level reconstruction agrees with need for R_{sep}
Now try ALPEN W + 3 parton event

SISCones solution including both partons (looking at inverted 2-D Snowmass potential)

2 partons clustered together
$\Delta R=0.8, z=0.4$

2-D projection of lego plot
Same ALPGEN (+PYTHIA) event at topocluster level
The LHC will be a very jetty place

- Total cross sections for t\(\bar{t}\) and Higgs production saturated by t\(\bar{t}\) (Higgs) + jet production for jet \(p_T\) values of order 10-20 GeV/c
- \(\sigma_{W+3\,\text{jets}} > \sigma_{W+2\,\text{jets}}\)

![Graph showing prediction for the production of W+ \(\geq 1, 2, 3\) jets at the LHC shown as a function of the transverse energy of the lead jet. A cut of 20 GeV has been placed on the other jets in the prediction.](image1)

- Indication that can expect interesting events at LHC to be very jetty (especially from gg initial states)
- Also can be understood from point-of-view of Sudakov form factors

![Graph showing the dependence of the LO t\(\bar{t}\)+jet cross section on the jet-defining parameter \(p_T,\text{min}\), together with the top pair production cross sections at LO and NLO.](image2)

![Graph showing the dependence of the LO t\(\bar{t}\)+jet cross section on the jet-defining parameter \(p_T,\text{min}\), together with the top pair production cross sections at LO and NLO.](image3)
Example: 4 jet event
ATLAS jet reconstruction

- Using calibrated topoclusters, ATLAS has a chance to use jets in a dynamic manner, not possible in any previous hadron-hadron calorimeter, i.e. to examine the impact of multiple jet algorithms/parameters/jet substructure on every event.

blobs of energy in the calorimeter correspond to 1/few particles (photons, electrons, hadrons); can be corrected back to hadron level rather than jet itself being corrected similar to running at hadron level in Monte Carlos.
UE/pileup corrections: Jet areas

determined by clustering ghost particles of vanishing energy; see jet references

note that the $k_T$ algorithm has the largest jet areas, SISCones the smallest and anti-$k_T$ the most regular; one of the reasons we like the antikt
Jet areas in presence of pileup

- Single W+4jets event, all matched to partons.
- SIScone and kT show decreased area in presence of pileup.

pileup nibbles away at perimeter of jet
Area-based correction: Cacciari/Salam/Soyez

1) Find low $p_T$ jets in event. ($< 10$ GeV) We use kT5jet.
2) From these, find average/median $p_T$ density of event $\rho$
3) Determine area $A$ of signal jets
4) Subtract “pileup/UE” estimate

$$p_{Tcor} = p_T - \rho A$$

• Black points used to find $p_T$ density
• Red points are then corrected according to Jet area

See presentations of Brian Martin in ATLAS jet meetings. Used in SpartyJet.
Aside: Photon isolation at the LHC

- From a theoretical perspective, it’s best to apply a *Frixione-style* isolation criterion, in which the amount of energy allowed depends on the distance from the photon; this has the advantage of removing the fragmentation contribution for photon production, as well as discriminating against backgrounds from jet fragmentation.
- But most of the energy in an isolation cone is from underlying event/pileup.
- At Les Houches, we started to develop (being continued by Mike Hance, Brian,…in ATLAS):
  - (1) an implementation of the Frixione isolation appropriate for segmented calorimeters.
  - (2) a hybrid technique that separates the UE/pileup energy from fragmentation contributions (using SpartyJet).

**Action Items:**
- Susan, Joey, Kajari, Jean-Philippe

**Exp:**
Look again in detail at the Frixione criterium, what is the impact at LHC of UE/PU, of fragmentation; see if some “hybrid” (simple cone vs Frixione) can be found, suitable for exp. application.

**Theory:**
use existing (and possibly upgraded) codes to study difference in x-sections obtained with Frixione-criterium and some “pedestal” allowed in the central cone.

Look also at “democratic” approach.
Jet masses

- Very useful if looking for resonance in boosted jet (top jet)
- Naturally produced by QCD radiation
- Depends on jet algorithm/size

In NLO pert theory

\[ \sqrt{p_{J\mu} p_{J}^{\mu}} = \sqrt{\left\langle M^2 \right\rangle_{NLO}} = f \left( \frac{p_{J}}{\sqrt{s}} \right) \sqrt{\alpha_s(p_{J})} (p_{J} R) \]

Rule-of-thumb

\[ \sqrt{\left\langle M^2 \right\rangle_{NLO}} \sim 0.2 p_{J} R \]

Fig. 53. The average jet mass is plotted versus the transverse momentum of the jet using several different jet algorithms with a distance scale \( D = R_{\text{conic}} \) of 0.7.
Distribution of jet masses

- Sudakov suppression for low jet masses
- fall-off as $1/m^2$ due to hard gluon emission
- algorithm suppression at high masses
  - jet algorithms tend to split high mass jets in two

Fig. 51. The jet mass distributions for an inclusive jet sample generated for the LHC with a $p_{T,\text{min}}$ value for the hard scattering of approximately 2 TeV/c, using several different jet algorithms with a distance scale ($D = R_{\text{cone}}$) of 0.7.
SpartyJet

J. Huston, K. Geerlings, Brian Martin
Michigan State University

P-A. Delsart, Grenoble

C. Vermillion, Washington

http://projects.hepforge.org/spartyjet/

If interested for ATLAS, please contact
Brian.thomas.martin@cern.ch
Summary

- We have an opportunity (forced on us) to understand the QCD environment at the LHC before we reach discover-potential integrated luminosities.
- We have the ability (with the ATLAS detector) to make more detailed measurements of final states including jets than any collider detector previous.
- Allons!
• Workhorse for many predictions at the LHC are still LO PDFs
• Many LO predictions at the LHC differ significantly from NLO predictions, not because of the matrix elements but because of the PDFs
• $W^+$ rapidity distribution is the poster child
  ◦ the forward-backward peaking obtained at LO is an artifact
  ◦ large $x_u$ quark distribution is higher at LO than NLO due to deficiencies in the LO matrix elements for DIS

Figure 1. A comparison of the NLO pseudodata for SM boson rapidity distributions (in $\Delta y=0.4$ bins) predicted at the LHC (14 TeV) to the respective LO predictions based on CTEQ6.6M and CTEQ6L1 PDFs.
Where are the differences between LO and NLO partons?

low $x$ and high $x$ for up

everywhere for gluon
Talking points

- LO* pdf’s should behave as LO as $x \rightarrow 0$; as close to NLO as possible as $x \rightarrow 1$
- LO* pdf’s should describe underlying event at Tevatron with a tune similar to CTEQ6L (for convenience) and extrapolate to a reasonable UE at the LHC
Modified LO PDFs

- Try to make up for the deficiencies of LO PDFs by:
  - relaxing the momentum sum rule
  - including NLO pseudo-data in the LO fit to guide the modified LO distributions
- Results tend to be in better agreement with NLO predictions, both in magnitude and in shape
- Some might say that the PDFs then have no predictive power, but this is true for any LO PDFs

- See arXiv:0910.4183; PDFs available from LHAPDF
- See arXiv:0711.2473 for MRST2007lomod PDFs

Figure 6. Predictions for the $W^+$ rapidity distribution at the LHC ($\sqrt{s} = 7, 10$ and $14$ TeV) in $\Delta y = 0.4$ bins, given at NLO using the CTEQ6.6M PDFs, and at LO using the CT09MC2 and MRST2007lomod PDFs. The actual cross sections (without normalization rescaling factors) are shown.
Higgs K-factor is too large to absorb into PDFs (nor would you want to)

Shape is ok with LO PDF’s, improves a bit with the modified LO PDFs
Now to NLO

- Less sensitivity to unphysical input scales, i.e. renormalization and factorization scales
- First level of prediction where normalization (and sometimes shape) can be taken seriously
- More physics
  - parton merging gives structure in jets
  - initial state radiation
  - more species of incoming partons
- Suppose I have a cross section $\sigma$ calculated to NLO ($O(\alpha_s^n)$)
- Any remaining scale dependence is of one order higher ($O(\alpha_s^{n+1})$)
  - in fact, we know the scale dependent part of the $O(\alpha_s^{n+1})$ cross section before we perform the complete calculation, since the scale-dependent terms are explicit at the previous order

\[
\frac{d\sigma}{dE_T} = \alpha_s(\mu_R)^2 A + \alpha_s(\mu_R)^3 (B + 2b_0 L A) + \alpha_s(\mu_R)^4 (C + 3b_0 L B + (3b_0^2 L^2 + 2b_1 L) A)
\]

with $L = \log(\mu_R/E_T)$ and $b_i$ the known beta function coefficients.

Inclusive jet prod at NNLO

we know $A$ and $B$, not $C$

LO has monotonic scale dependence, non-monotonic at NLO

Figure 11: Single jet inclusive distribution at $E_T = 100$ GeV and $0.1 < |\eta| < 0.7$ at $\sqrt{s} = 1800$

The NNLO coefficient $C$ is unknown. The curves show the guesses $C = 0$ (solid) and $C = \pm B^2/A$ (dashed).
Remember the smallest uncertainty

...confirmed by NNPDF

Z produced at high $p_T$ (~200 GeV/c) is produced primarily from gq scattering in this x range can we use Z production at high $p_T$ as a normalization?

Errors (%) on $tt/Z(+j)$ ratios using CTEQ6ME

<table>
<thead>
<tr>
<th>$\frac{\delta \sigma(tt)}{\sigma(Z)}$</th>
<th>$\frac{\delta \sigma(tt)}{\sigma(Z+j)}$</th>
<th>$\frac{\delta \sigma(tt)}{\sigma(Z+j, p_T &gt; 50)}$</th>
<th>$\frac{\delta \sigma(tt)}{\sigma(Z+j, p_T &gt; 100)}$</th>
<th>$\frac{\delta \sigma(tt)}{\sigma(Z+j, p_T &gt; 150)}$</th>
<th>$\frac{\delta \sigma(tt)}{\sigma(Z+j, p_T &gt; 200)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>6.3</td>
<td>5.7</td>
<td>5.2</td>
<td>4.3</td>
<td>3.5</td>
</tr>
</tbody>
</table>

Idea appears to work!
tT pdf error vs Z (+ jet)

% pdf uncertainty

% pdf uncertainty

Z (% pdf uncertainty)
tT pdf error vs Z (+ jet)

p_{T} > 25 \text{ GeV/c}
tT pdf error vs Z (+ jet)

- tT vs Zj50
tT pdf error vs Z (+ jet)
for $p_{Z_T} > 200$ GeV/c, not only does the $Z$ uncertainty become small, but there's also a correlation developing with the $tT$ cross section, because of the gluon being in a similar $x$ range.
What do the eigenvectors mean?

- Each eigenvector corresponds to a linear combination of all 20 (22,24) pdf parameters, so in general each eigenvector doesn’t mean anything?

- However, with 20 (22,24,26) dimensions, often eigenvectors will have a large component from a particular direction

- Take eigenvector 1 (for CTEQ6.1); error pdf’s 1 and 2

- It has a large component sensitive to the small x behavior of the u quark valence distribution

- Not surprising since this is the best determined direction
What do the eigenvectors *mean*?

- Take eigenvector 8 (for CTEQ6.1); error pdf’s 15 and 16
- No particular direction stands out
What do the eigenvectors mean?

- Take eigenvector 15 (for CTEQ6.1); error pdf’s 29 and 30
- Probes high x gluon distribution

creates largest uncertainty for high $p_T$ jet cross sections at both the Tevatron and LHC

Figure 29. The pdf errors for the CDF inclusive jet cross section in Run 1 for the 20 different eigenvector directions. The vertical axes show the fractional deviation from the central prediction and the horizontal axes the jet transverse momentum in GeV.
For some events, the jet structure is very clear and there’s little ambiguity about the assignment of towers to the jet.

But for other events, there is ambiguity and the jet algorithm must make decisions that impact precision measurements.

If comparison is to hadron-level Monte Carlo, then hope is that the Monte Carlo will reproduce all of the physics present in the data and influence of jet algorithms can be understood.

- more difficulty when comparing to parton level calculations