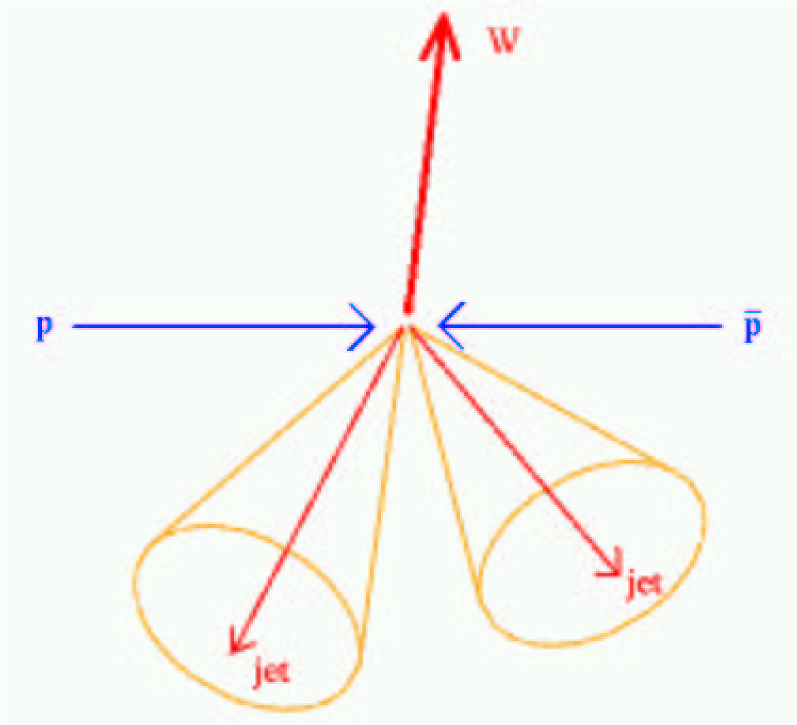
Boson + jet production using the Monte Carlo MCFM

Keith Ellis
FNAL

In collaboration with:
John Campbell (code and phenomenology)
and Dave Rainwater (phenomenology)

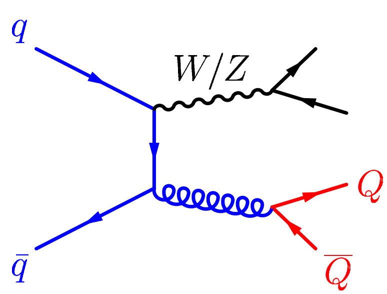
W +2 jet events

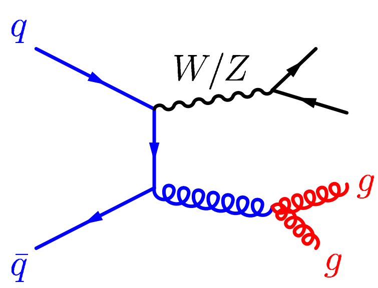
- Many such events at Run I of the Tevatron. For example, with an integrated luminosity of $108\ \mathrm{pb}^{-1}$ CDF collected $51400\ W \to e\nu$ events, of which 2000 are $W + 2$ jet events. This yields an $80\mathrm{pb}$ cross-section.



W+2 jet theory

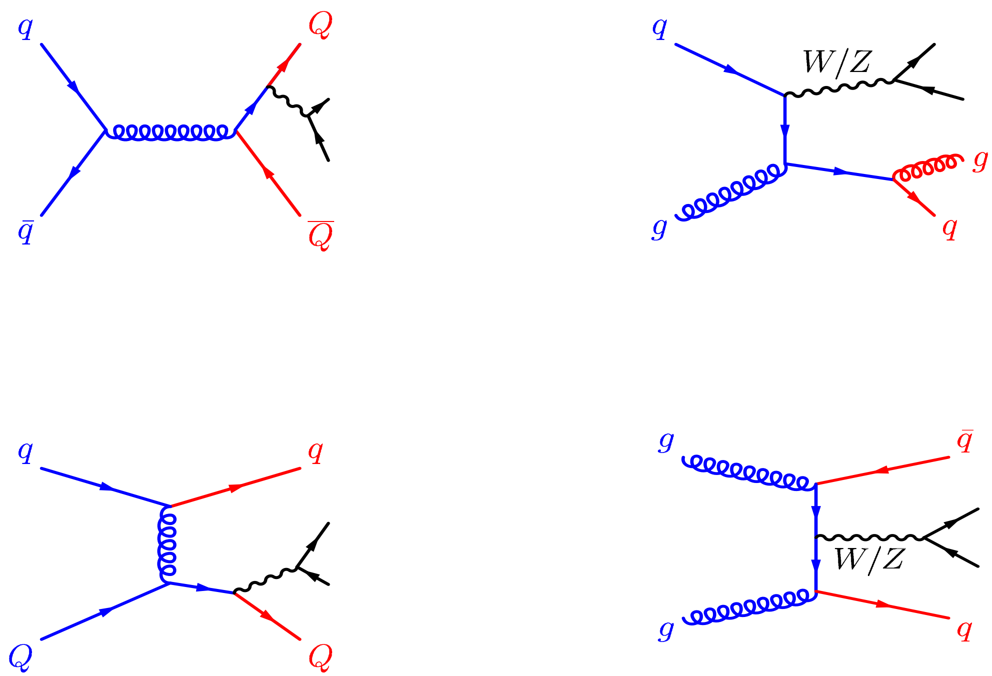
- In the leading order of perturbative QCD, this process can be represented by Feynman tree-graphs.
- At leading order a jet is represented by a single final state quark or gluon (Local Parton-Hadron Duality).
- There are two classes of diagrams at leading order, 4 quark and 2 quark, 2 gluon.





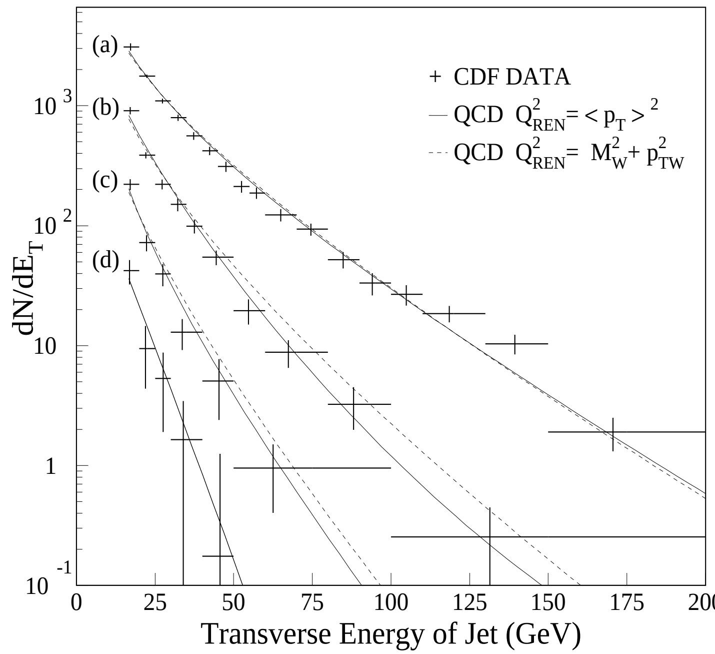
W +2 *jet theory, continued*

- Related diagrams provide other initial states that also contribute:



Multi-jet data

- This theory describes multi-jet data fairly well. For example, the leading-jet E_T spectrum for $W + n$ jet production ($n = 1, \ldots, 4$):



- Deficiency at high E_T in the $W + 1$ jet sample.

Failings of leading order

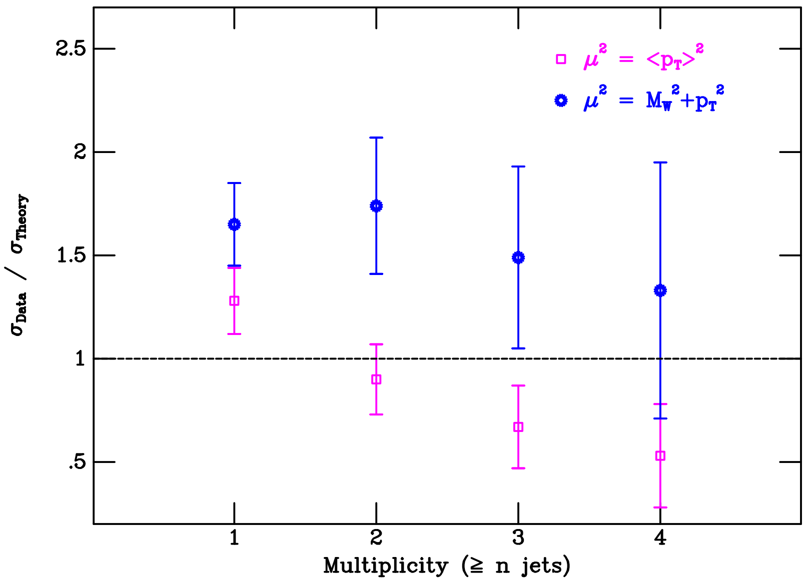
- Some discrepancies arise when the theory is examined in more detail.
- An important theoretical input is the value of the renormalization and factorization scales, μ_R and μ_F.
- These artificial variables are required only because we cannot solve the full theory of QCD. Instead, we compute an observable $\mathcal{O}_{\rm full}$ perturbatively,

$$\mathcal{O}_{\rm full}^{\rm W+2\ jet} = \alpha_S^2\, \mathcal{O}_2 + \alpha_S^3\, \mathcal{O}_3 + \ldots + \alpha_S^r\, \mathcal{O}_r + \ldots$$

- Truncating this series produces a dependence upon μ_R and μ_F in our predictions.
- In the leading order picture $= \mathcal{O}_2$.

Scale worries

- $W+ \geq n$ jets cross-sections from CDF Run I, compared with (enhanced) leading order theory:

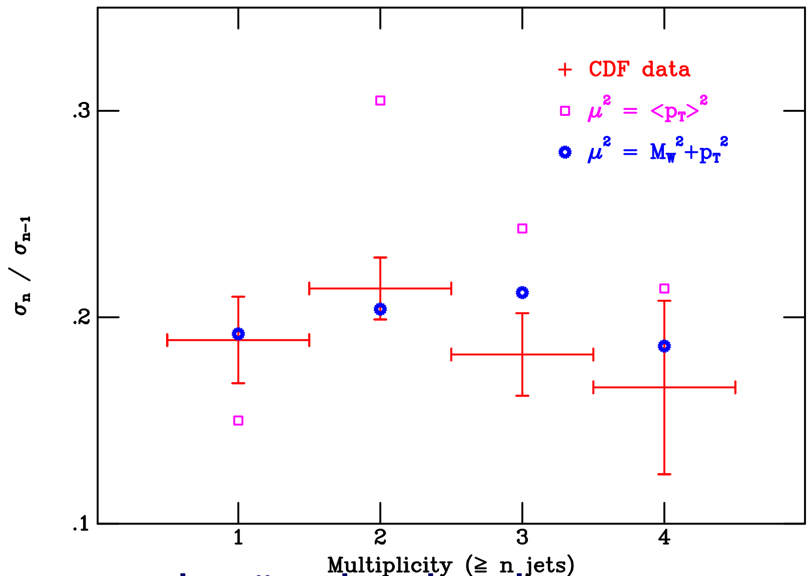


$$\mu_R = \mu_F \equiv \mu$$

- To reproduce the raw cross-sections, especially for the $W + 1$, 2 jet data, the low scale $\mu^2 = \langle p_T \rangle^2$ is preferred.

Scale worries, continued

- Ratio of n-jet cross sections, σ_n/σ_{n-1}:



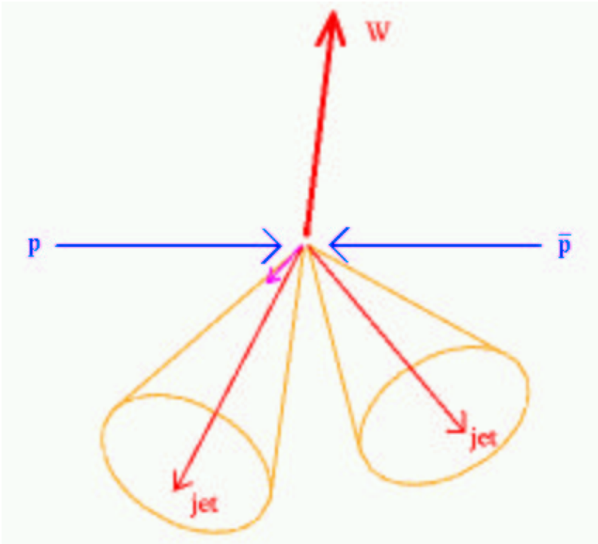
$$\mu_R = \mu_F \equiv \mu$$

- Measures the "reduction in cross section caused by adding a jet" (roughly $\sim \alpha_S$).
- Useful quantity since systematics should cancel.
- High scale $\mu^2 = M_W^2 + p_T^2$ now much closer to data.

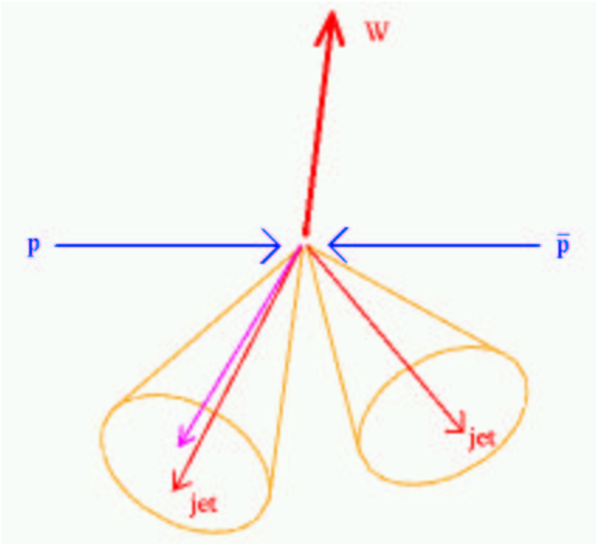
Next-to-leading order

- At next-to-leading order, we include an extra “unresolved” parton in the final state

soft



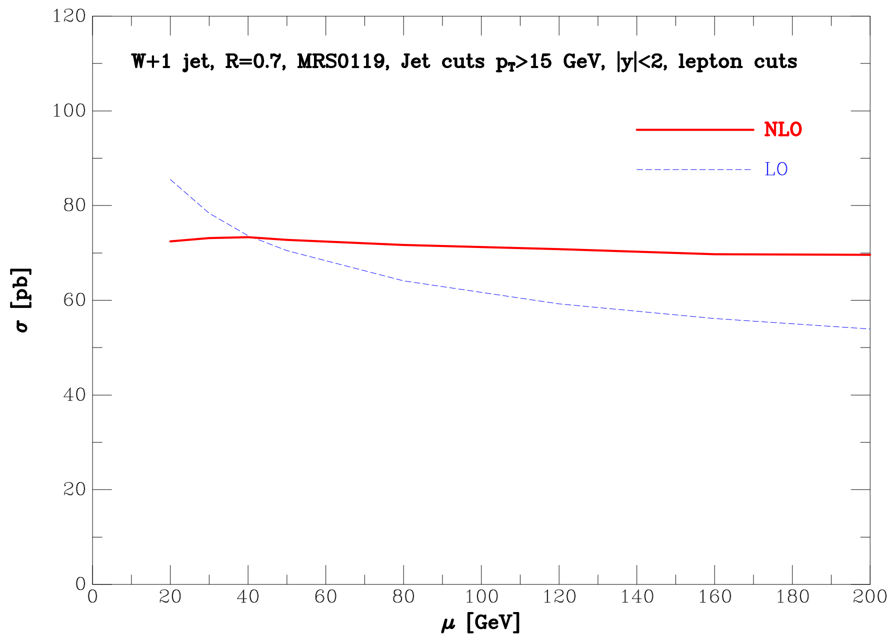
collinear



- The theory begins to look more like an experimental jet, so one expects a better agreement with data.

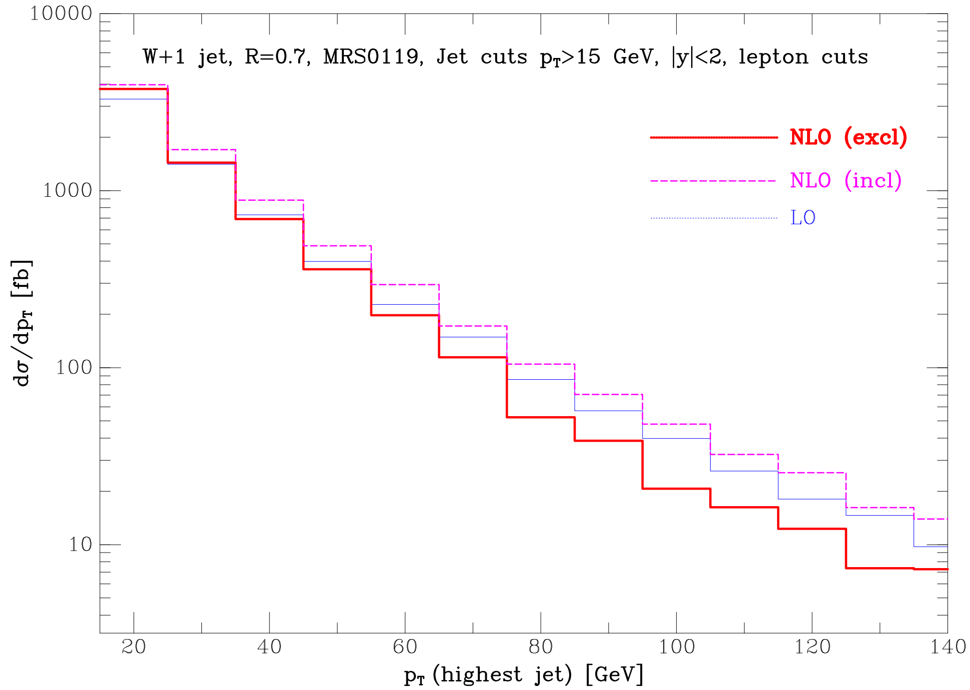
Scale dependence

- $W + 1$ jet cross-section demonstrates the reduced scale dependence that is expected at NLO, as large logarithms are partially cancelled.



- Change between low ~ 20 GeV and high ~ 80 GeV scales is about 30% at LO and $< 5\%$ at NLO.

Jet p_T distribution



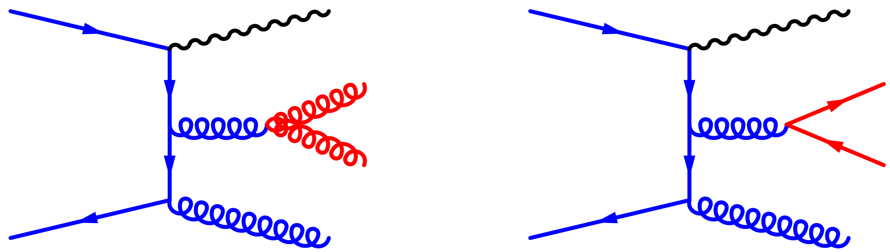
$\mu = 80$ GeV

- Leading E_T jet becomes much softer at NLO.
- Exclusive $\rightarrow$ depletion at high-E_T, since jets there are more likely to radiate a parton passing the jet cuts
  Inclusive $\rightarrow$ shape more similar to LO

$W + 2$ *jets, NLO theory*

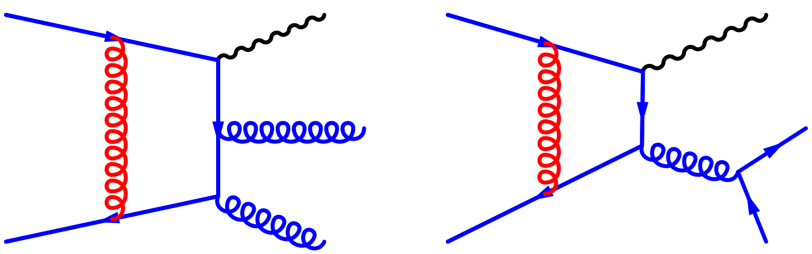
- Feynman diagrams for extra parton radiation, e.g.

soft gluon



collinear quarks

- Loop diagrams, also one extra factor of α_S:

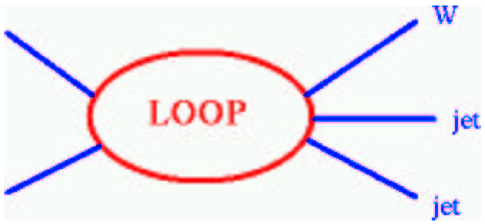


NLO difficulties

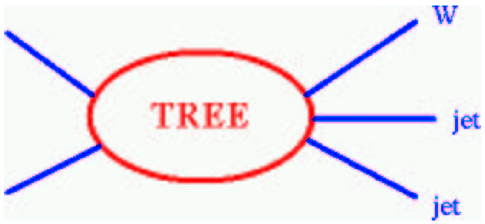
- We must somehow combine two types of diagrams, each with a different number of final state partons.
- Whilst this procedure is well understood from the theory point of view, it does raise problems:
  - There is no simple correspondence between a data event and the theory description.
  - Interfacing with Pythia,Herwig is difficult, since one must be careful not to double-count soft and collinear radiation. However, there has been some progress in this area recently for relatively simple processes.
  - Less experimental familiarity with NLO generators.

Loop diagrams

- Use the helicity amplitudes of Z. Bern et al.
- Loop integrals are divergent. The usual choice is to regularize in $d = 4 - 2\epsilon$ dimensions.
- Simplistically, the result is:



$$= \left(\frac{A}{\epsilon^2} + \frac{B}{\epsilon} + C\right) \times$$

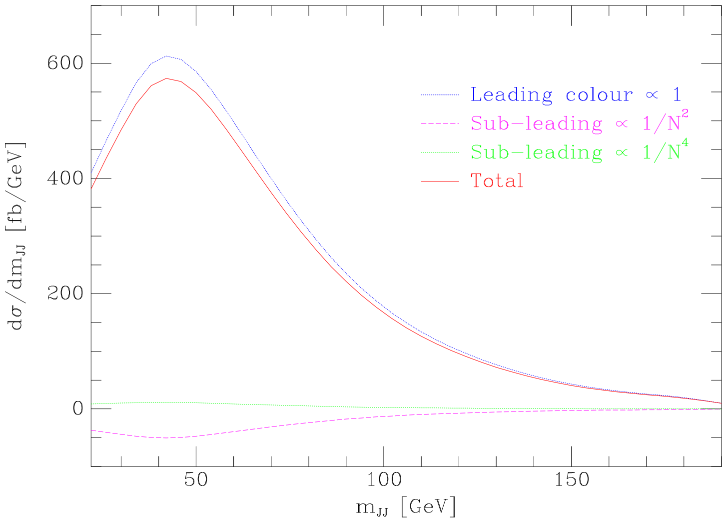


+ finite terms

- The finite terms are rational functions of the invariants, log's and di-log's. There are many terms and they are also slow to evaluate.
- Calculation is organized using a colour decomposition.

Colour decomposition

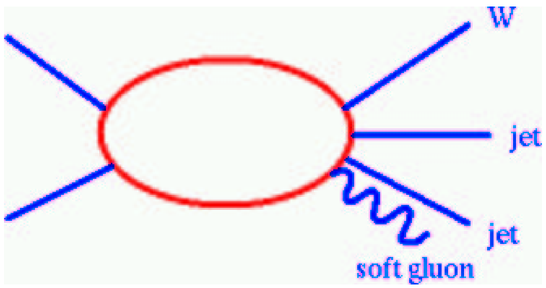
- Recall the two classes of diagrams - ones involving 2 quarks, 2 gluons and those with 4 quarks. We can write the matrix elements for these diagrams as an expansion in the number of colours, N.
- The 2 quark, 2 gluon diagrams contain the leading term and pieces suppressed by $1/N^2$ and $1/N^4$. The 4 quark diagrams are suppressed by $1/N$ and $1/N^3$.



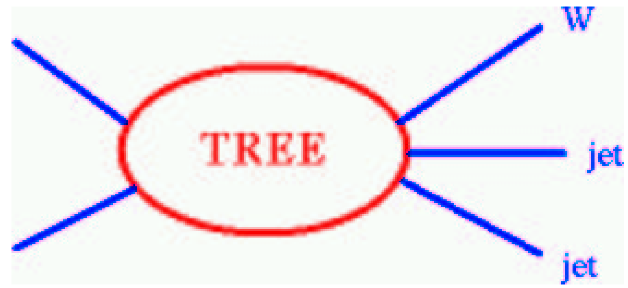
dijet mass distribution

Real diagrams

- The matrix elements for the production of $W+2$ jets with an extra soft gluon are also divergent, for example in the limit $E_{gluon} \to 0$.
- the (colour-ordered) matrix elements factorize:



$\longrightarrow$ eikonal $\times$ factor



- The eikonal factor contains all the soft singularities.
- By partial fractioning one can apportion this into two terms which have different collinear singularities.

$$\frac{p_i \cdot p_j}{p_i \cdot k \; p_j \cdot k} = \frac{p_i \cdot p_j}{(p_i \cdot k + p_j \cdot k) \; p_i \cdot k} + (i \leftrightarrow j)$$

- Exploit this to construct the counterterms.

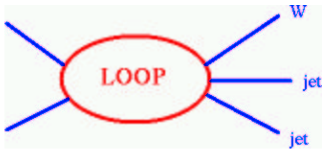
Real diagrams, continued

- Now we must compensate for the singularities that we just cancelled.
- This is done by analytically integrating the eikonal factor over the phase space of the soft gluon, to give:

$$\int (\text{eikonal factor})\, dPS = \frac{D}{\epsilon^2} + \frac{E}{\epsilon} + F$$

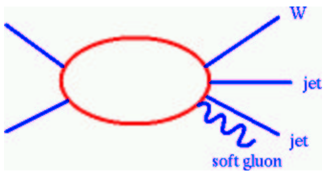
- This is called the subtraction method.
- Careful choice of the kinematics in the lowest-order matrix elements is made, to optimize the singularity cancellation - the dipole subtraction scheme.

Result



$$= \left(\frac{A}{\epsilon^2} + \frac{B}{\epsilon} + C\right) \times$$

$$\int$$



$$dPS^{\,\mathrm{gluon}} = \left(\frac{D}{\epsilon^2} + \frac{E}{\epsilon} + F\right) \times$$

- $A = -D$, $B = -E$ . . . so all poles cancel (KLN).
- We are left with integrals over the final 2-jet phase-space for:
  - The remaining finite parts of the loop diagrams;
  - The non-singular real emission diagrams where one jet contains a soft gluon or a collinear quark.

$W+2$ *jet outline*

1. Assemble all loop matrix elements.
2. Assemble all real radiation matrix elements.
3. Enumerate all possible soft, collinear singularities.
4. Construct appropriate counterterms to cancel these.
5. Check the cancellation occurs in the singular limits.
6. Integrate over the singular areas of phase-space.
7. Check that these poles cancel with those from loops.
8. With a given jet definition and cuts, perform the phase-space integration.
9. Accumulate predictions for any observables required.

MCFM Summary - v. 3.4

$p\bar{p} \to W^\pm/Z$	$p\bar{p} \to W^+ + W^-$
$p\bar{p} \to W^\pm + Z$	$p\bar{p} \to Z + Z$
$p\bar{p} \to W^\pm + \gamma$	$p\bar{p} \to W^\pm/Z + H$
$p\bar{p} \to W^\pm + g^\star\ (\to b\bar{b})$	$p\bar{p} \to Zb\bar{b}$
$p\bar{p} \to W^\pm/Z +$ 1 jet	$p\bar{p} \to W^\pm/Z +$ 2 jets
$p\bar{p}(gg) \to H$	$p\bar{p}(gg) \to H +$ 1 jet
$p\bar{p}(VV) \to H +$ 2 jets	

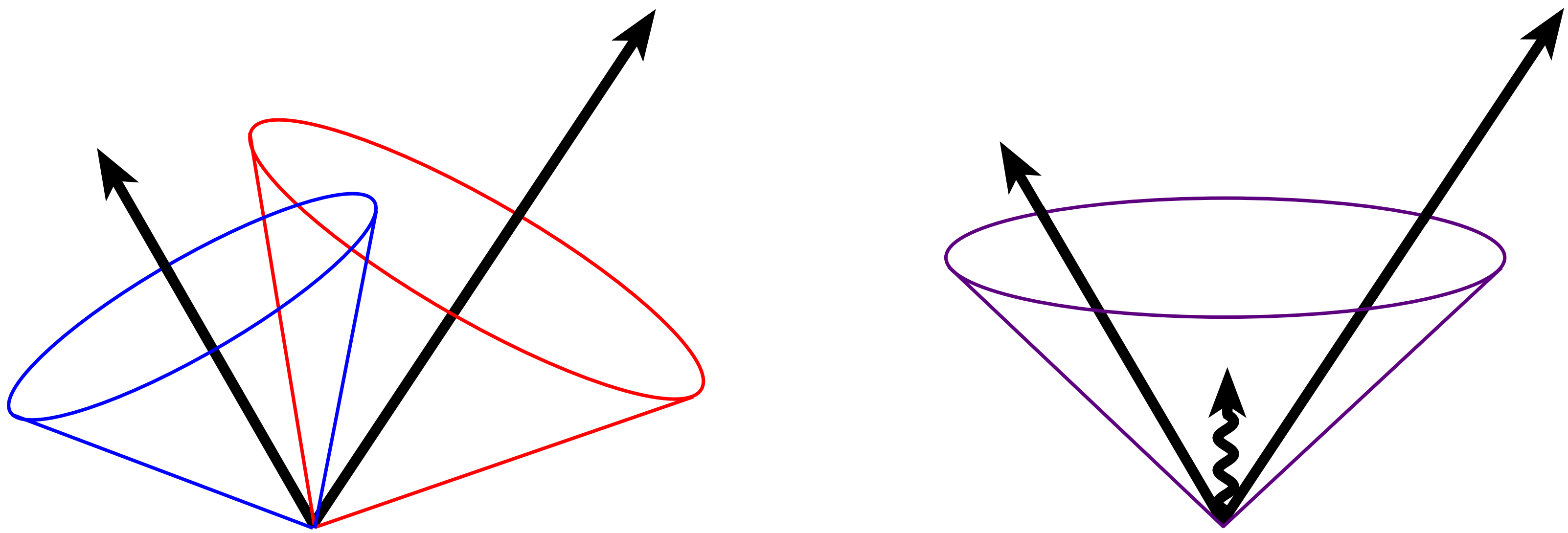
- MCFM aims to provide a unified description of a number of processes at NLO accuracy.
- Various leptonic and/or hadronic decays of the bosons are included as further sub-processes.
- MCFM version 2.0 is part of the CDF code repository.

MCFM Information

- Version 3.4 available at:
  `http://mcfm.fnal.gov`
- Improvements over previous release:
  - more processes
  - better user interface
  - support for PDFLIB, Les Houches PDF accord
  - ntuples as well as histograms
  - unweighted events
  - Pythia/Les Houches generator interface (LO)
- Coming attractions:
  - even more processes, photon fragmentation etc.

Defining a jet - cone algorithm

- Cone-based algorithm, $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} > R$.
- Very popular in Run I.
- Suffers from sensitivity to soft radiation at NLO.



- Instability can be mitigated by extra jet seeds, e.g. midpoint algorithms.

Defining a jet - k_T algorithm

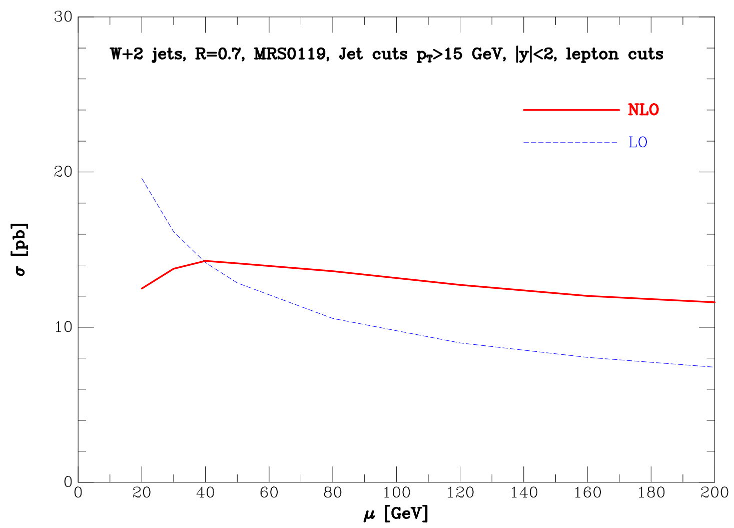
- Preferred by theory - insensitive to soft radiation, immediate matching to resummed calculations.
- Limited experimental use at hadron colliders due to difficulties with energy subtraction.
- Jets are clustered according to the relative transverse momentum of one jet with respect to another.
- Similarity with cone jets is kept, since the algorithm still terminates with all jets having $\Delta R > R$.
- We shall adopt the k_T prescription that is laid out for Run II (G. Blazey et al.), where other ambiguities such as the jet recombination scheme are fixed.

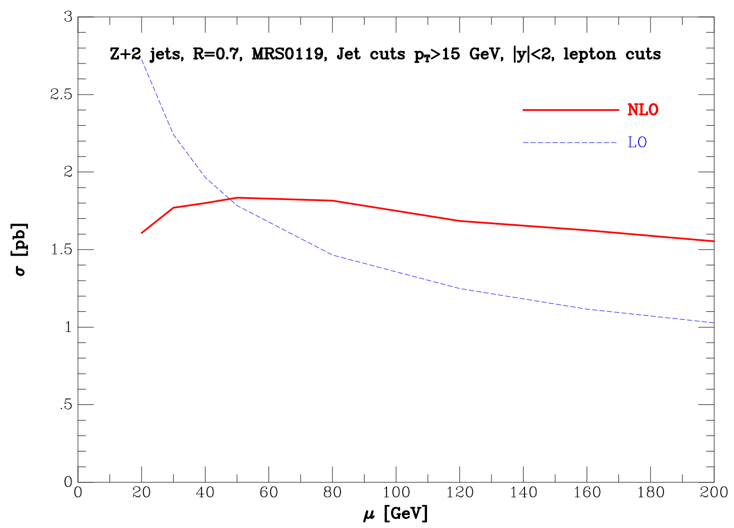
Tevatron event cuts

- k_T clustering algorithm with pseudo-cone size, $R = 0.7$.
- Jet cuts:
  $p_T^{\rm jet} > 15$ GeV, $|y^{\rm jet}| < 2$.
- Exclusive cross-section - so exactly 2 jets.
- Lepton cuts:
  $p_T^{\rm lepton} > 20$ GeV, $|y^{\rm lepton}| < 1$.
- (W only) Missing transverse momentum:
  $p_T^{\rm miss} > 20$ GeV.
- (Z only) Dilepton mass:
  $m_{e^-e^+} > 15$ GeV (since γ^* is also included).

Scale dependence

- Choose equal factorization and renormalization scales.
- Examine scale dependence of the cross-section integrated over 20 GeV $< m_{JJ} < 200$ GeV.

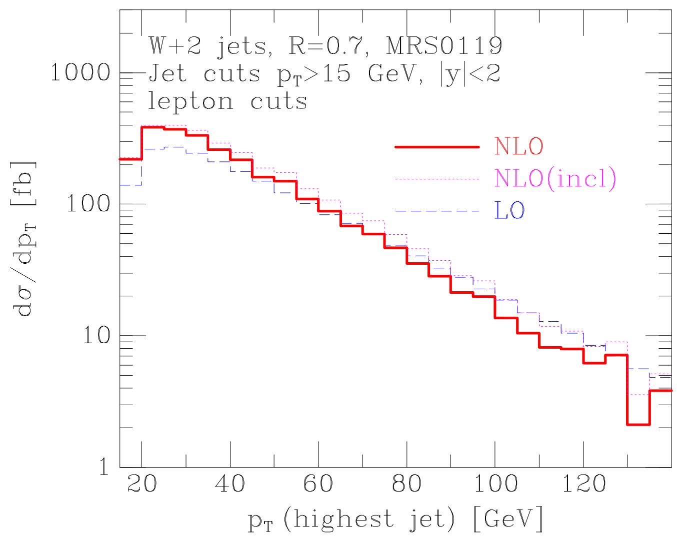


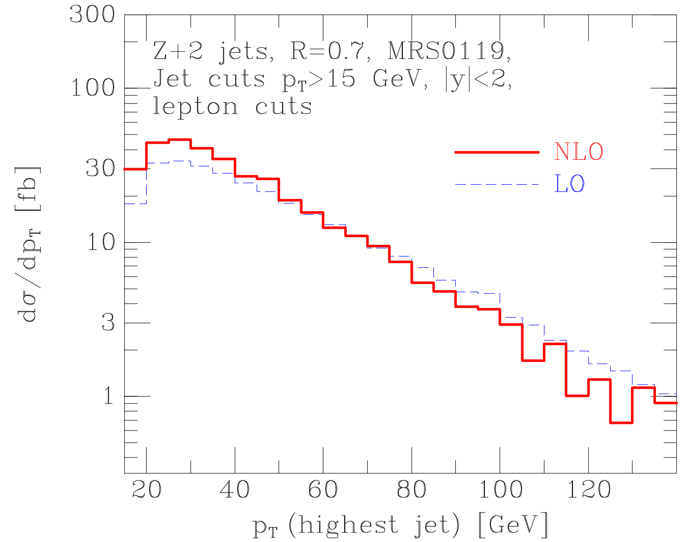


- Scale dependence much reduced from $\sim 100\%$ to $\sim 10\%$ in both cases.

Leading p_T distribution

- p_T distribution of the hardest jet in $W, Z{+}2$ jet events, at the scale $\mu = 80$ GeV.

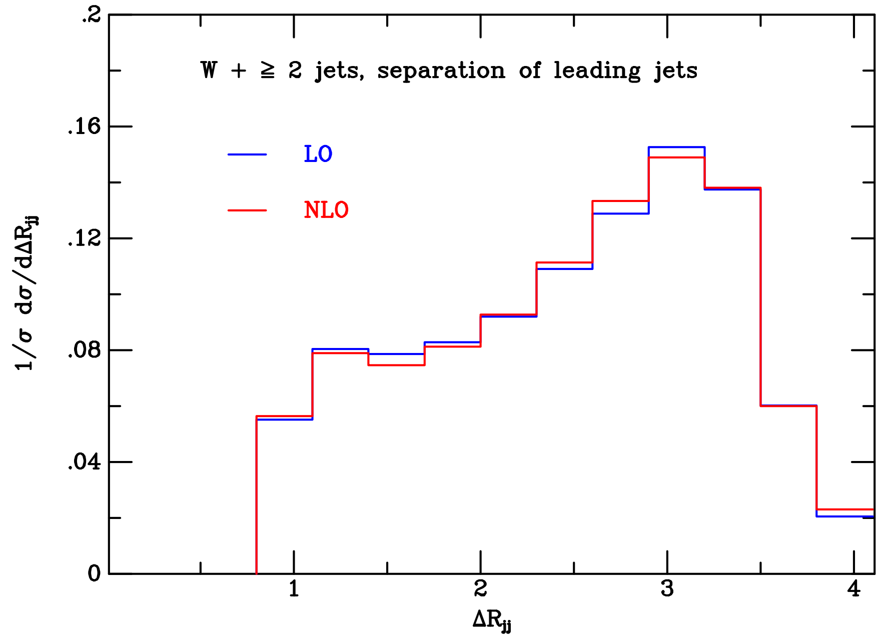




- Turn-over at low p_T since 15 GeV $< p_T^2 < p_T^1$.
- The exclusive spectrum is much softer at next-to-leading order, as in the 1-jet case.
- High-E_T tail is ‘filled in’ for the inclusive case.

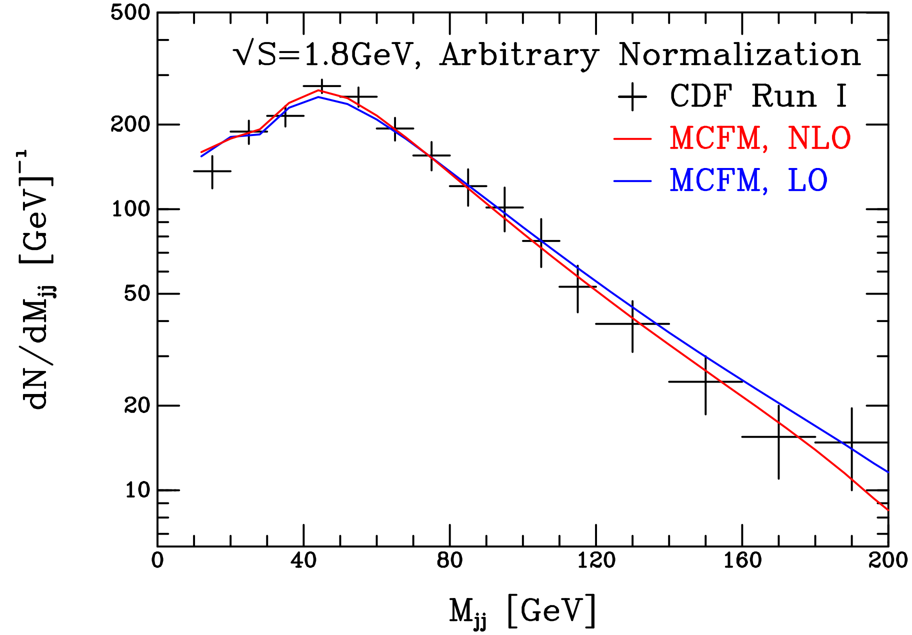
Jet-jet separation

- In Run I, there was some discrepancy in the shape of the jet-jet separation ΔR_{jj} compared with LO theory.
- Results at NLO appear to confirm the leading order shape, with no significant dependence on scale.



Di-jet mass in Run I

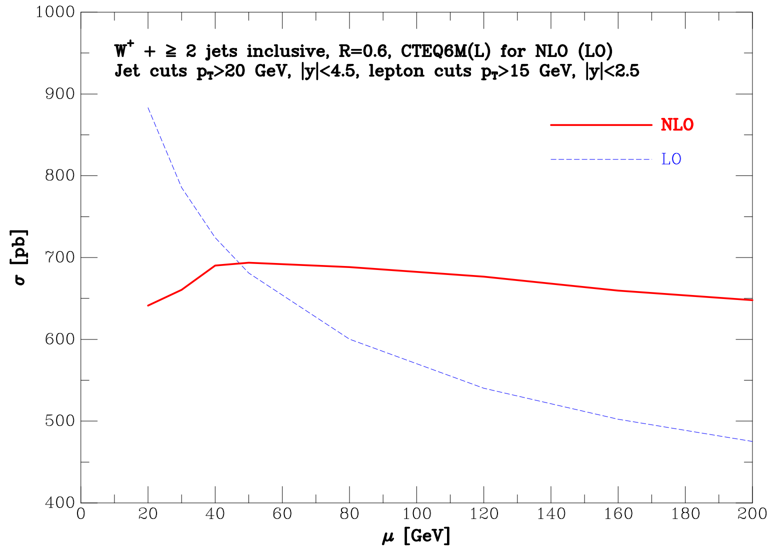
- For the W+2 jet inclusive cross-section, compare the predicted dijet mass distribution with data, allowing the total cross-section to float.

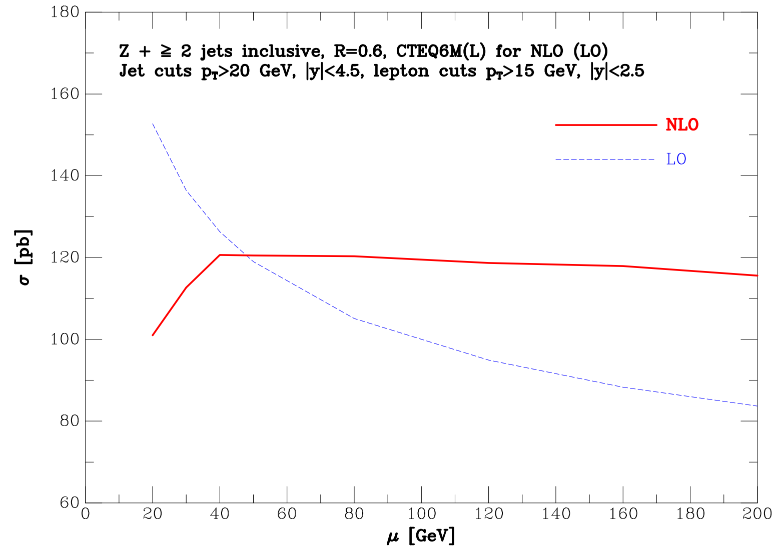


- Much better agreement with NLO result, especially towards both ends of the distribution.

$V + \geq 2$ *jets at the LHC*

- Different set of cuts at $\sqrt{s} = 14$ TeV and here we consider the inclusive cross section.

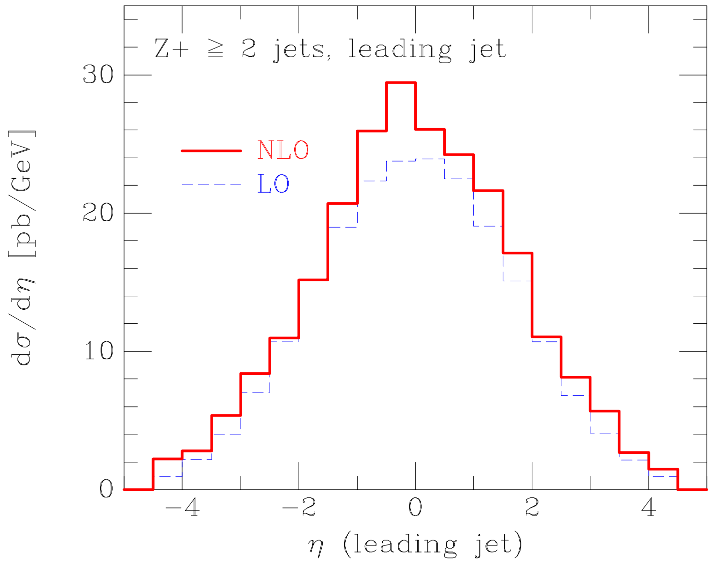


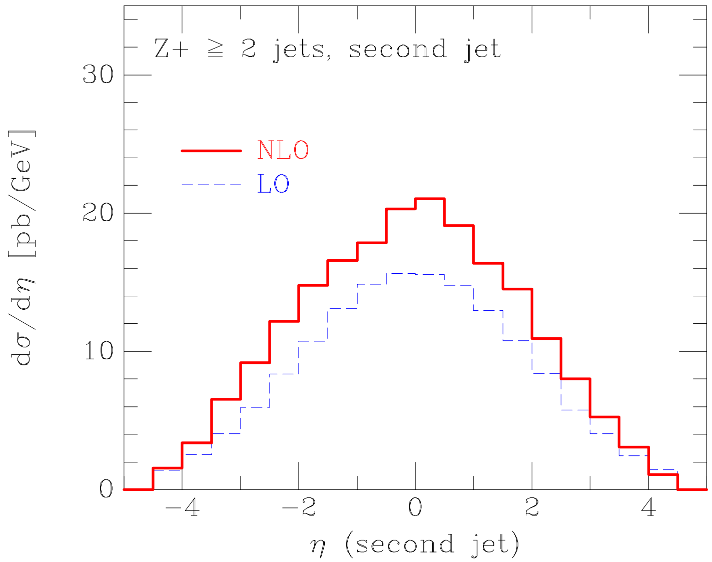


- The NLO corrections are somewhat smaller than at the Tevatron, approximately $10 - 20\%$ around $\mu = M_W$. Much less sensitivity to the scale μ.

Jet rapidities at the LHC

- The shapes of the jet rapidity distributions do not change significantly at next-to-leading order.





- Further study of these processes at the LHC is underway.

Heavy flavour content

- Many signals of new physics involve the production of a W or Z boson in association with a heavy particle that predominantly decays into a $b\bar{b}$ pair.
- A light Higgs is a prime example and will provide a promising search channel in Run II.

$$p\bar{p} \longrightarrow W(\to e\nu)H(\to b\bar{b})$$

$$p\bar{p} \longrightarrow Z(\to \nu\bar{\nu}, \ell\bar{\ell})H(\to b\bar{b})$$

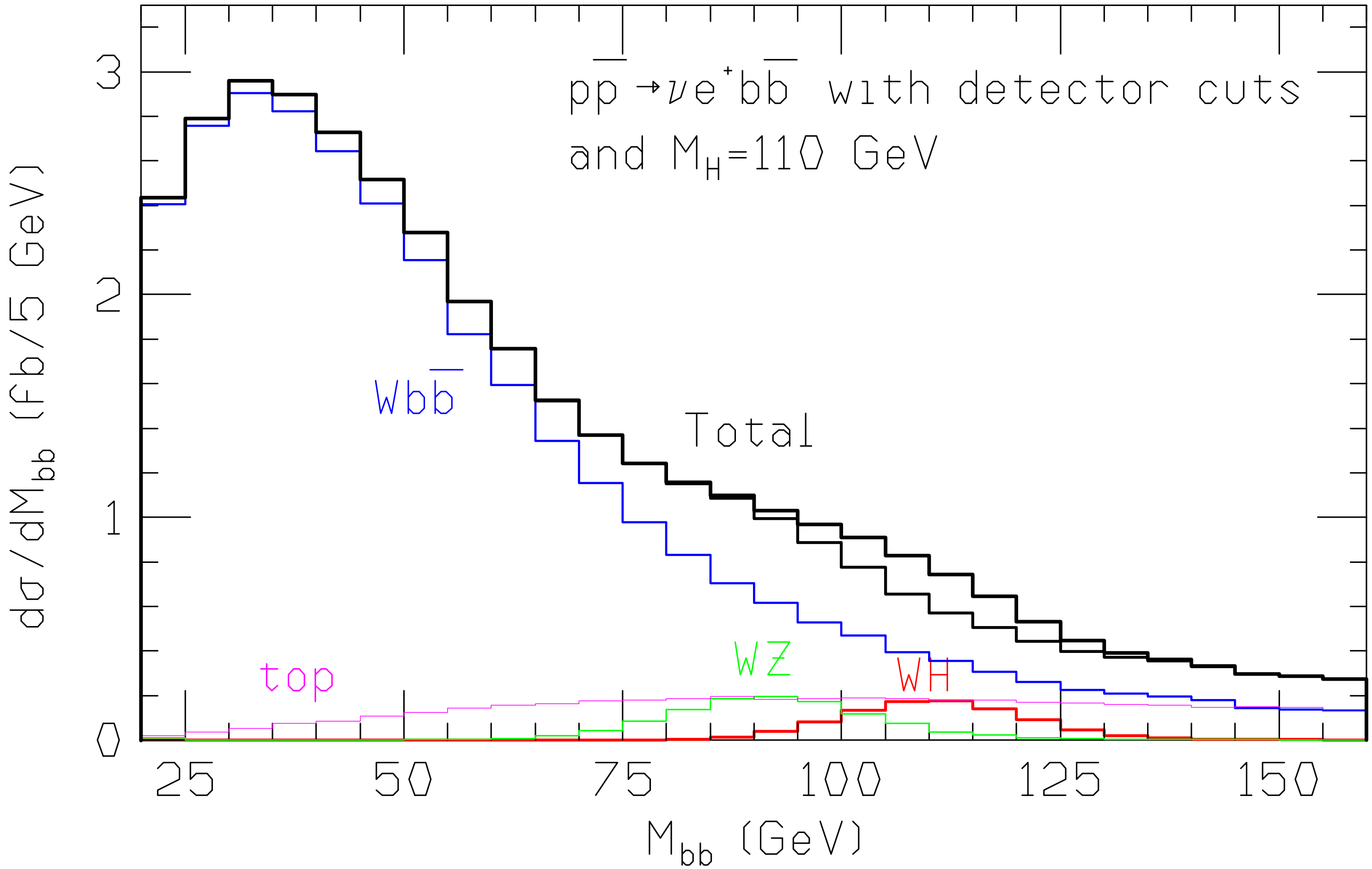
- However, we will need to understand our SM backgrounds very well to perform this search.
- The largest background is ‘direct’ production:

$$p\bar{p} \longrightarrow W\, g^{\star}(\to b\bar{b})$$

$$p\bar{p} \longrightarrow Z\, b\bar{b}$$

Background importance

- NLO study of WH search using MCFM.



Predicting the $Wb\bar{b}$ background

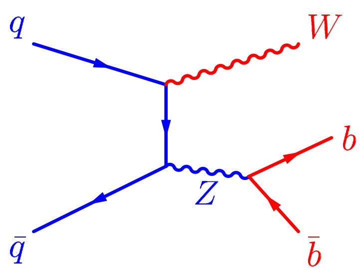
- There are a number of methods for predicting the Standard Model 'direct' background.
- Amongst the theoretical choices are:
  - Fixed order vs. event generator;
  - LO vs. NLO;
  - Pythia vs. Herwig;
  - Massive b's vs. Massless b's.
- Citing a 40% uncertainty on the leading-order calculation (M. Mangano), a recent study by CDF uses a mixed approach relying heavily on generic $W+$ jet data, but with some theoretical input.

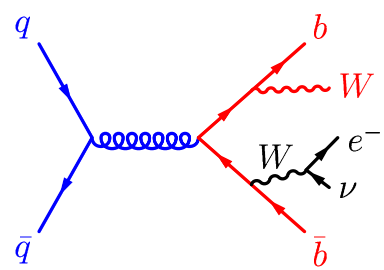
Hybrid recipe

1. Measure the number of $W + 2$ jet events.
2. Subtract the number of events predicted by theory from non-direct channels.
   - $t\bar{t}$ (Pythia norm. to NLO)
   - Diboson (Pythia norm. to NLO)
   - Single top (Pythia/Herwig norm. to NLO)
3. This estimates the number of direct $W + 2$ jet events.
4. Use VECBOS (leading order) + Herwig to estimate the fraction of $W + 2$ jet events that contain two b's.
5. Obtain prediction for direct $W + b\bar{b}$ events.

Other $Wb\bar{b}$ backgrounds

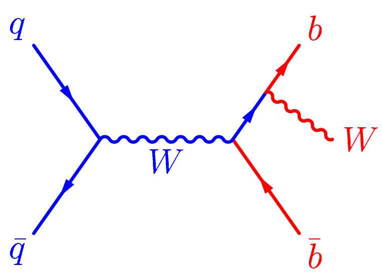
diboson

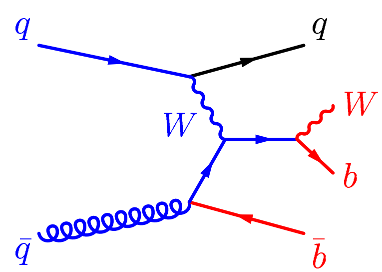




$t\bar{t}$

single
top (s)





single
top (t)

Alternatives

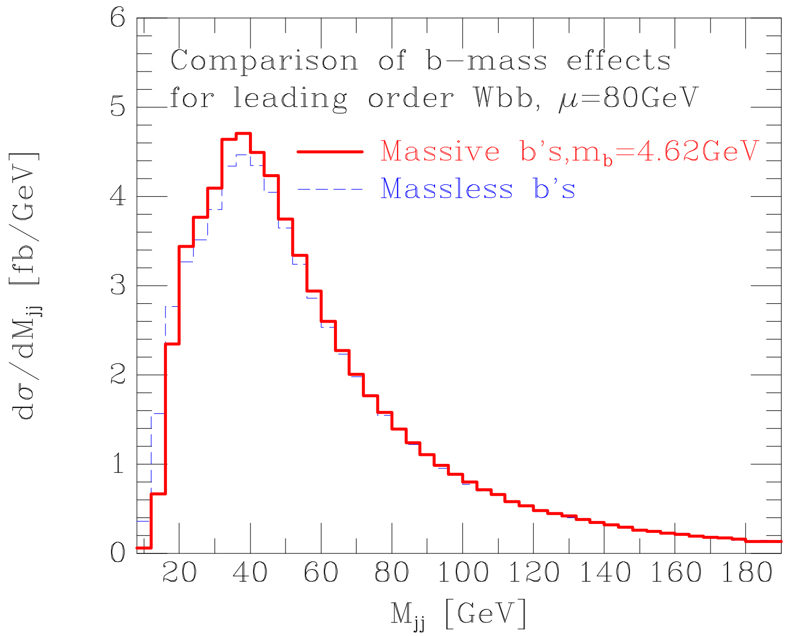
- Is this the best we can do?
- VECBOS suffers from the same leading order uncertainty that we were trying to avoid.
- We can calculate the $Wb\bar{b}$ cross-section at NLO in MCFM. This has a much reduced scale dependence, but suffers from no showering and massless b's.
- Another option is to calculate the same fraction,

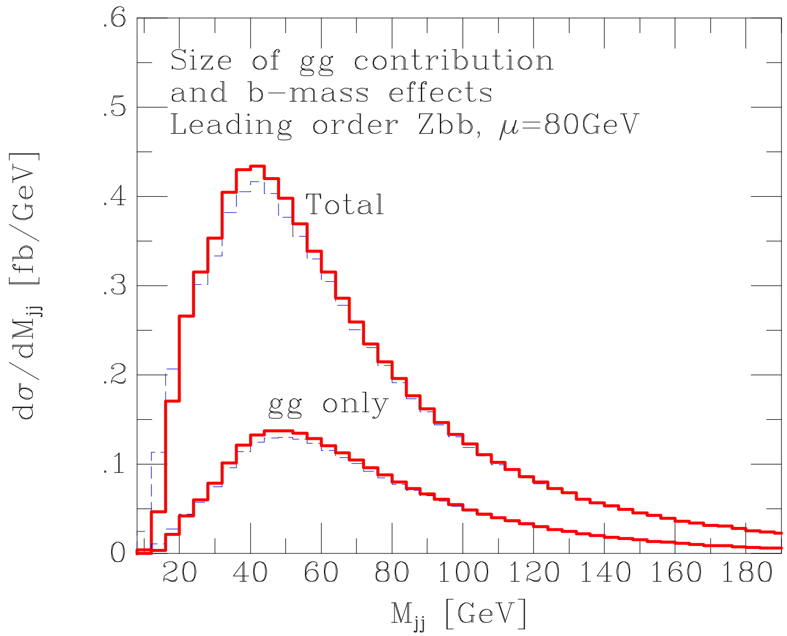
$$\frac{\sigma(Wb\bar{b})}{\sigma(W+2\,\mathrm{jet})}$$

that is calculated by Herwig, but at NLO. Some systematics (showering, perhaps) should cancel.

b-mass effects

- Compare the lowest order predictions for m_b zero and non-zero.

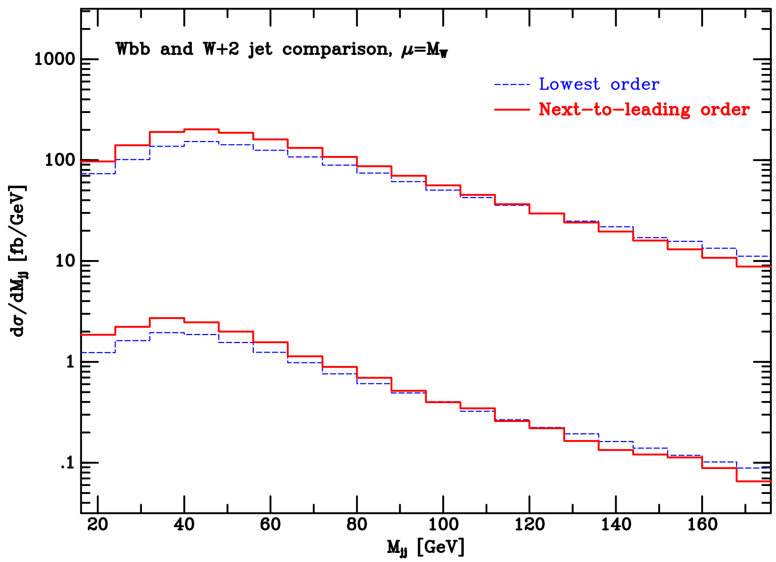


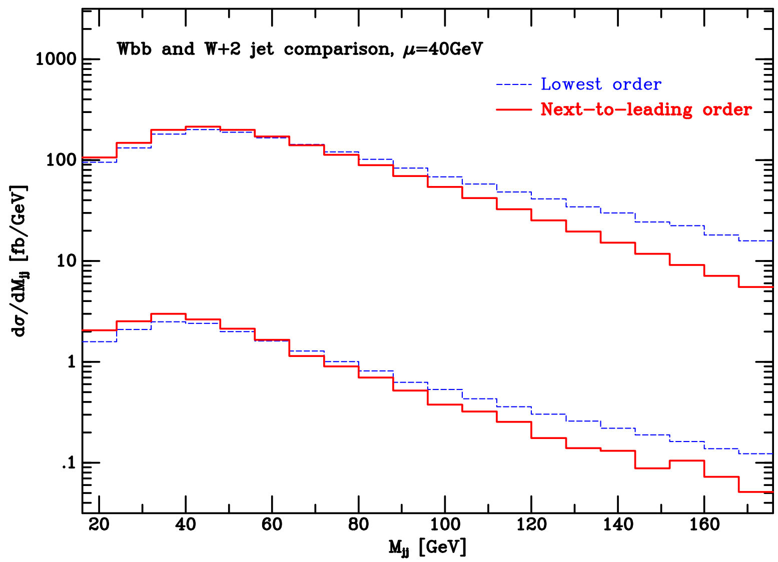


- In the interesting region - the peak at low mass - matrix element effects dominate over phase space. The corrections there are of order 5%.

m_{JJ} distributions

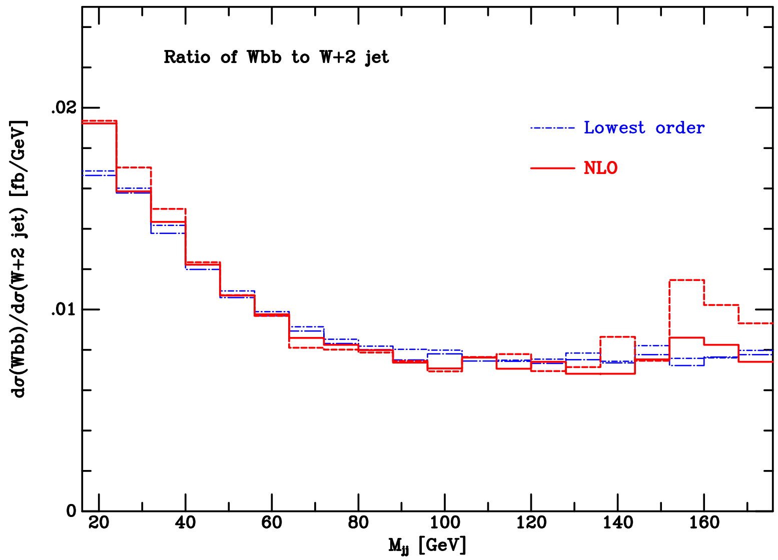
- $Wb\bar{b}$ and $W+2$ jet distributions appear very similar in shape at both LO and NLO. The shapes change when moving to a lower scale, with a depletion in the cross-section at high M_{jj}.





Heavy flavour fraction

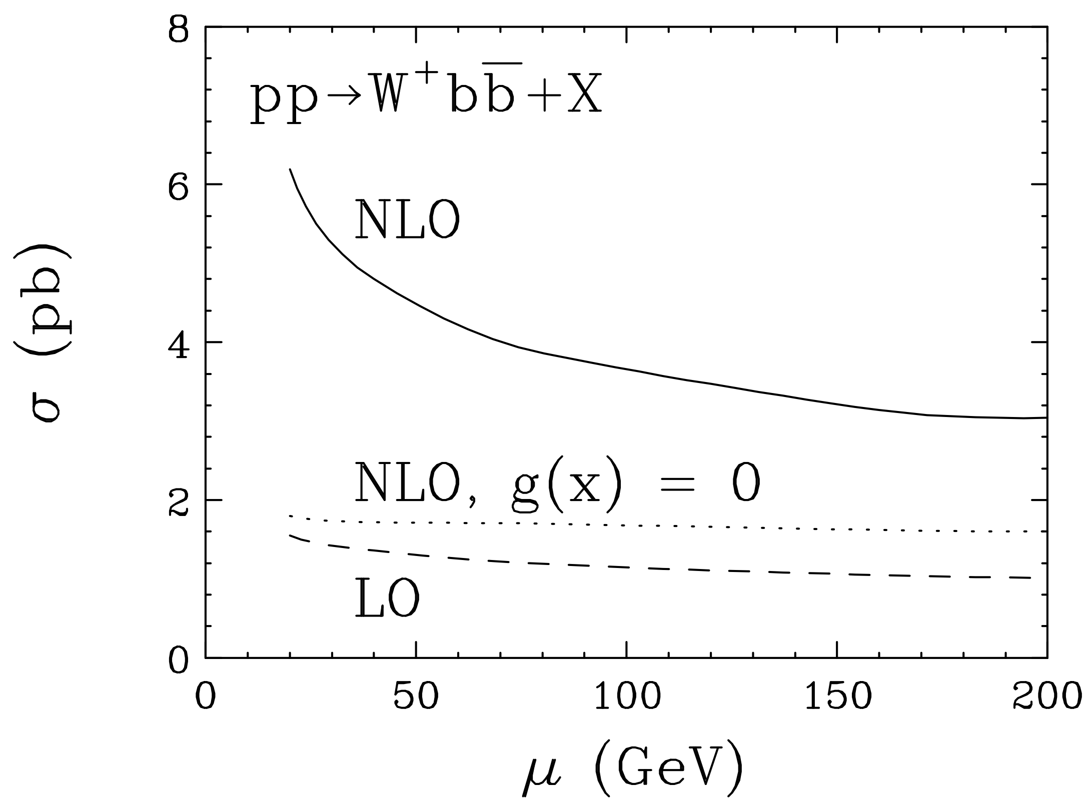
- The ratio of b-tagged to untagged jets changes very little at NLO and appears to be predicted very well by perturbation theory.



- The fraction is peaked at low M_{jj}, where it is approximately 2.5 times as high as the fairly constant value of 0.8% for $M_{jj} > 60$ GeV.

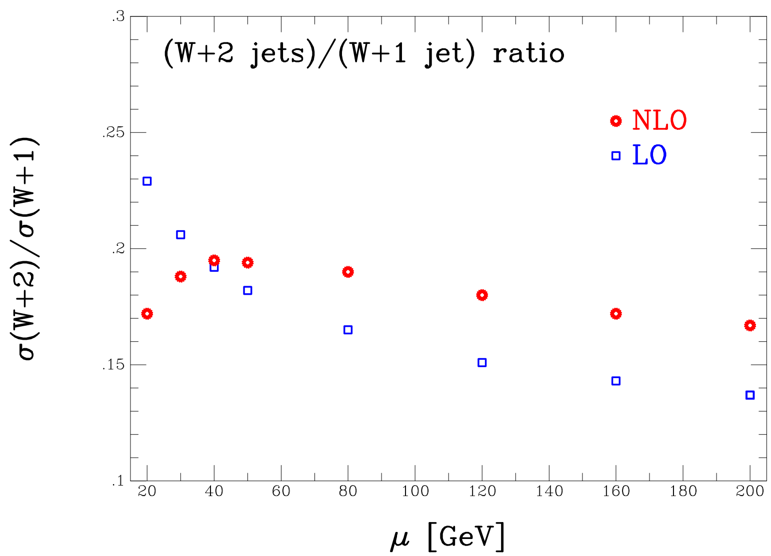
$W b\bar{b}$ at LHC

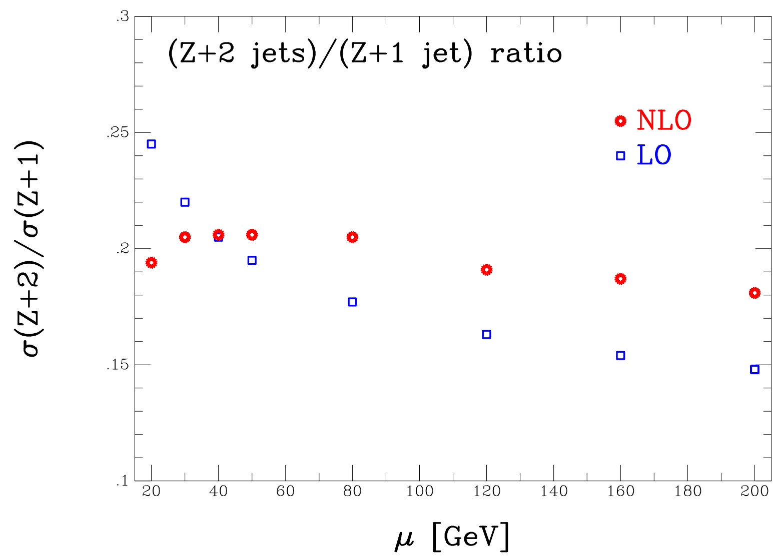
- Large corrections to $W b\bar{b}$ rate at the LHC
- NLO introduces a gluon-initiated process which is not present at LO



Cross-section ratios, revisited

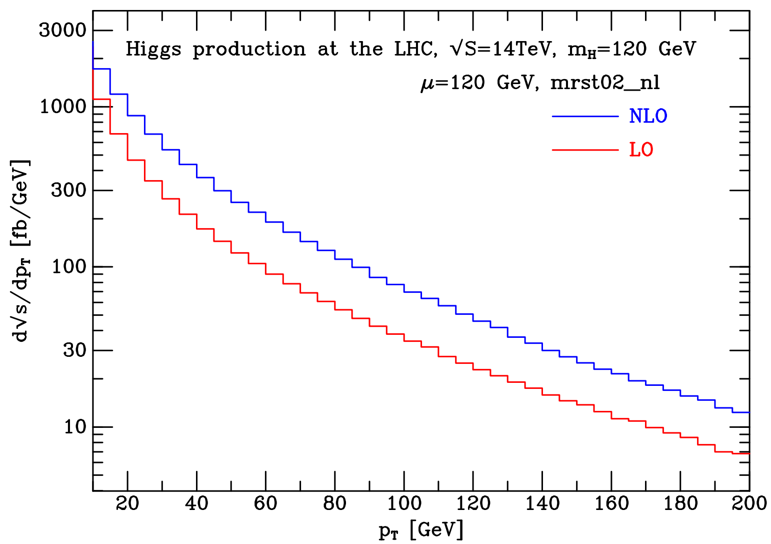
- Ratios calculated for Run II, so not directly comparable to previous data.

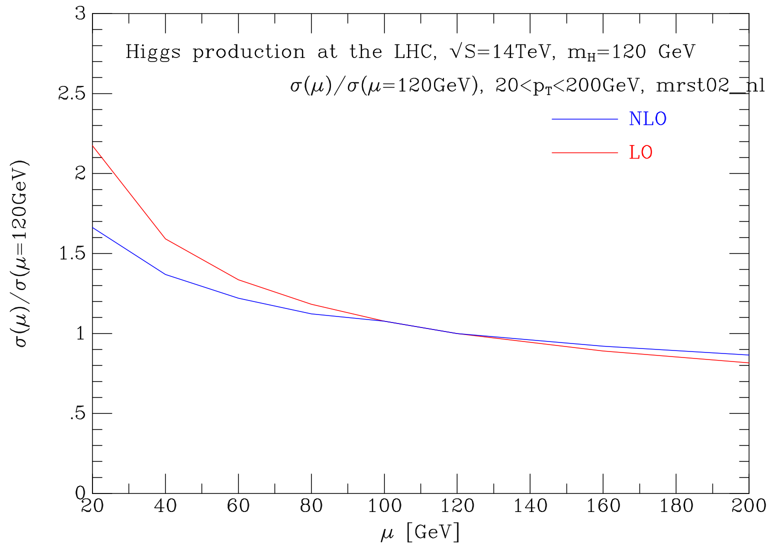




- Much more stable at NLO than LO, particularly in the region of conventional scales $\sim 30 - 80$ GeV.
- Precision measurements of α_S in the near future?.

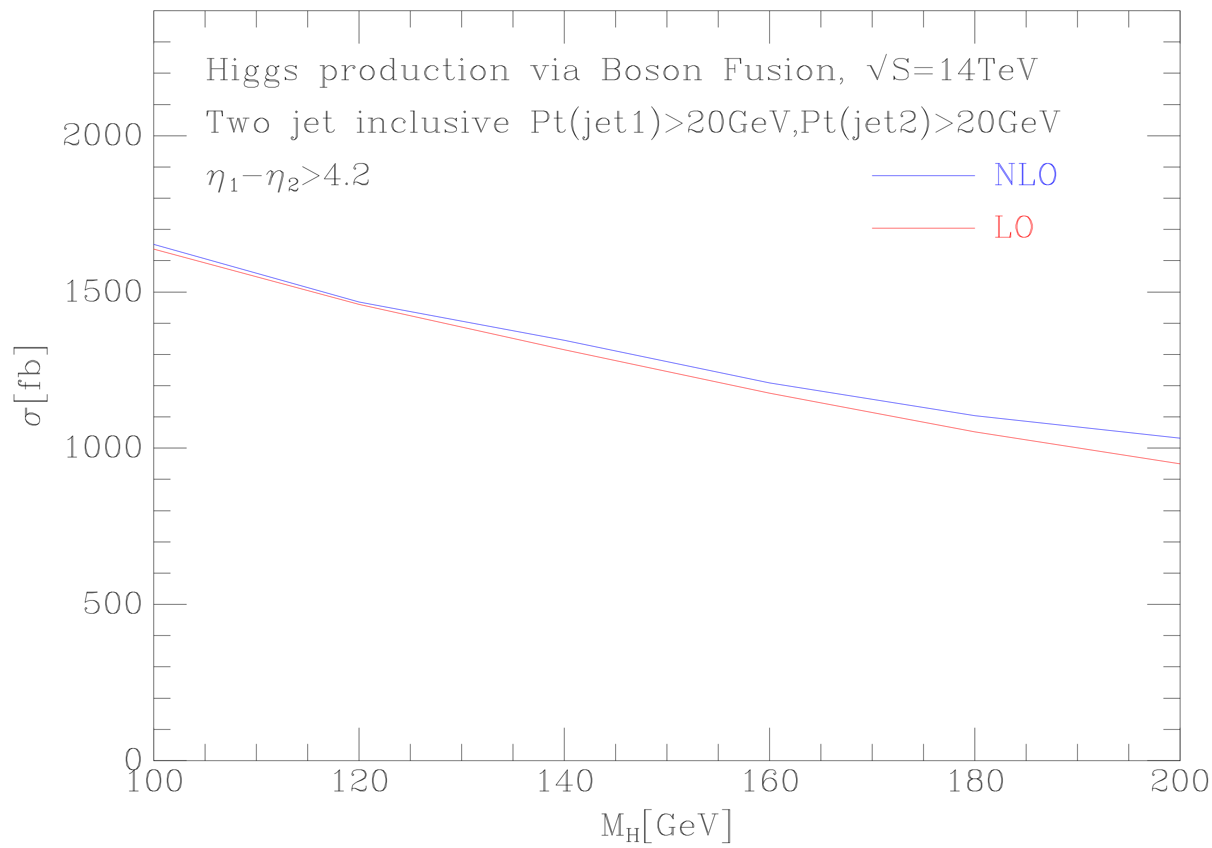
Higgs + one jet predictions





- Results at the LHC
- Calculation performed in $m_t \to \infty$ limit.
- Substantial change in normalization at NLO
- Somewhat more stable at NLO than LO.
- Previous calculations,

Higgs + two jet predictions



- Vector boson fusion calculation
- Small corrections, Han et al, PRL.69:3274,1992

Conclusions

- The NLO corrections for $W/Z+2$ jets have been calculated and are available at `mcfm.fnal.gov`.
- Scale dependence is greatly reduced to $\sim 10\%$ and distributions are considerably changed upon including QCD corrections.
- NLO code is contained in MCFM v3.4. Current code in the CDF repository is v2.0 and will be updated soon.
- The fraction of a $W+2$ jet sample that contains two b-jets is predicted very well in perturbation theory.
- There are many interesting studies to be done - from tests of QCD to backgrounds for new physics.