MCFM details and applications

October 2005

John Campbell
CERN

MCFM details

Overview

- Downloadable general purpose NLO code, "MCFM" JC and R.K. Ellis
(+F. Tramontano, +F. Maltoni, S. Willenbrock)

$p\bar{p} \to W^\pm/Z$	$p\bar{p} \to W^+ + W^-$
$p\bar{p} \to W^\pm + Z$	$p\bar{p} \to Z + Z$
$p\bar{p} \to W^\pm + \gamma$	$p\bar{p} \to W^\pm/Z + H$
$p\bar{p} \to W^\pm + g^\star\,(\to b\bar{b})$	$p\bar{p} \to Zb\bar{b}$
$p\bar{p} \to W^\pm/Z +$ 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	$p\bar{p} \to t + q$
$p\bar{p} \to H + b$	$p\bar{p} \to Z + b$
$p\bar{p} \to W + t$	

- NLO knowledge can be used in various ways.
  - ⋆ production of pairs of W's and Z's: the structure of the weak interaction at high energy
  - ⋆ $W/Z+$jets: testing QCD and sources of background events
  - ⋆ $H + 2$ jets: an important discovery mode at the LHC

Comparison with other approaches

- There are generic routines for handling common tasks, so that implementing new processes is “painless”
- Emphasis has been on bringing together calculations of signals and backgrounds for particularly challenging searches, so that NLO effects may be more easily studied with just one code
- Where possible, appropriate decays of vector bosons are included and all possible spin correlations are retained for a better assessment of the effect of experimental cuts
  - ⊖ low particle multiplicity (no showering)
  - ⊖ no hadronization
  - ⊖ hard to model detector effects
  - ⊕ less sensitivity to μ
  - ⊕ rates are better normalized
  - ⊕ fully differential distributions

MCFM Information

- Version 4.1 available at:

  `http://mcfm.fnal.gov`
- Improvements over previous releases:
  - ⋆ more processes
  - ⋆ better user control
  - ⋆ separate variation of factorization and renormalization scales
  - ⋆ support for PDFLIB, Les Houches PDF accord ($\longrightarrow$ PDF uncertainties)
  - ⋆ ntuples as well as histograms
  - ⋆ unweighted events
  - ⋆ basic event generator interface
  - ⋆ 'Behind-the-scenes' efficiency

most suitable for LO calculations

MCFM basics

- Written in Fortran 77.
- Operation of the program is mostly controlled by two files: `mdata.f` and `input.DAT`.
- The file `mdata.f` contains information on particle masses and couplings, for example:

```
data ewscheme  / -1                   /   ! Chooses EW scheme
data Gf_inp    / 1.16639d-5           /   ! G_F
data aemmz_inp / 7.7585538055706d-03 /   ! alpha_EM(m_Z)=1/128.89
data xw_inp    / 0.2312d0             /   ! sinˆ2(theta_W)
data wmass_inp / 80.419d0             /   ! W mass
data zmass_inp / 91.188d0             /   ! Z mass
```

- Most options for running the program are specified in `input.DAT`, with reference to the accompanying file `process.DAT`.
- The simplest output from the program is in the form of histograms (a standard set, but may be customized $\longrightarrow$ `nplotter.f`).

process.DAT

```
'4.2'           [file version number]
1  '  f(p1)+f(p2) --> W^+(-->nu(p3)+e^+(p4))'
6  '  f(p1)+f(p2) --> W^-(-->e^-(p3)+nu~(p4))'
1000 'nul'
11 '  f(p1)+f(p2) --> W^+(-->nu(p3)+e^+(p4))+f(p5)'
12 '  f(p1)+f(p2) --> W^+(-->nu(p3)+e^+(p4))+gamma(p5)'
13 '  f(p1)+f(p2) --> W^+(-->nu(p3)+e^+(p4))+c~(p5)'
14 '  f(p1)+f(p2) --> W^+(-->nu(p3)+e^+(p4))+c~(p5) [massless]'
16 '  f(p1)+f(p2) --> W^-(-->e^-(p3)+nu~(p4))+f(p5)'
17 '  f(p1)+f(p2) --> W^-(-->e^-(p3)+nu~(p4))+gamma(p5)'
18 '  f(p1)+f(p2) --> W^-(-->e^-(p3)+nu~(p4))+c(p5)'
19 '  f(p1)+f(p2) --> W^-(-->e^-(p3)+nu~(p4))+c(p5) [massless]'
1000 'nul'
20 '  f(p1)+f(p2) --> W^+(-->nu(p3)+e^+(p4)) +b(p5)+b~(p6) [massive]'
21 '  f(p1)+f(p2) --> W^+(-->nu(p3)+e^+(p4)) +b(p5)+b~(p6)'
22 '  f(p1)+f(p2) --> W^+(-->nu(p3)+e^+(p4)) +f(p5)+f(p6)'
23 '  f(p1)+f(p2) --> W^+(-->nu(p3)+e^+(p4)) +f(p5)+f(p6)+f(p7)'
24 '  f(p1)+f(p2) --> W^+(-->nu(p3)+e^+(p4)) +b(p5)+b~(p6)+f(p7)'
25 '  f(p1)+f(p2) --> W^-(-->e^-(p3)+nu~(p4)) +b(p5)+b~(p6) [massive]'
26 '  f(p1)+f(p2) --> W^-(-->e^-(p3)+nu~(p4)) +b(p5)+b~(p6)'
27 '  f(p1)+f(p2) --> W^-(-->e^-(p3)+nu~(p4)) +f(p5)+f(p6)'
28 '  f(p1)+f(p2) --> W^-(-->e^-(p3)+nu~(p4)) +f(p5)+f(p6)+f(p7)'
29 '  f(p1)+f(p2) --> W^-(-->e^-(p3)+nu~(p4)) +b(p5)+b~(p6)+f(p7)'
1000 'nul'
31 '  f(p1)+f(p2) --> Z^0(-->e^-(p3)+e^+(p4))'
32 '  f(p1)+f(p2) --> Z^0(-->3*(nu(p3)+nu~(p4)))'
33 '  f(p1)+f(p2) --> Z^0(-->b(p3)+b~(p4))'
1000 'nul'
41 '  f(p1)+f(p2) --> Z^0(-->e^-(p3)+e^+(p4))+f(p5)'
42 '  f(p1)+f(p2) --> Z_0(-->3*(nu(p3)+nu~(p4)))-[sum over 3 nu]+f(p5)'
43 '  f(p1)+f(p2) --> Z^0(-->b(p3)+b~(p4))+f(p5)'
1000 'nul'
44 '  f(p1)+f(p2) --> Z^0(-->e^-(p3)+e^+(p4))+f(p5)+f(p6)'
45 '  f(p1)+f(p2) --> Z^0(-->e^-(p3)+e^+(p4))+f(p5)+f(p6)+f(p7)'
1000 'nul'
48 '  f(p1)+f(p2) --> Z^0(-->e^-(p3)+e^+(p4))+gamma(p5)'
49 '  f(p1)+f(p2) --> Z^0(-->3*(nu(p3)+nu~(p4)))-[sum over 3 nu]+gamma(p5)'
1000 'nul'
50 '  f(p1)+f(p2) --> Z^0(-->e^-(p3)+e^+(p4))+b~(p5)+b(p6) [massive]'
51 '  f(p1)+f(p2) --> Z^0(-->e^-(p3)+e^+(p4))+b(p5)+b~(p6)'
52 '  f(p1)+f(p2) --> Z_0(-->3*(nu(p3)+nu~(p4)))+b(p5)+b~(p6)'
53 '  f(p1)+f(p2) --> Z^0(-->b(p3)+b~(p4))+b(p5)+b~(p6)'
56 '  f(p1)+f(p2) --> Z^0(-->e^-(p3)+e^+(p4))+b(p5)+b~(p6)+f(p7)'
```

- Each line corresponds to a process that can be calculated, mostly at next-to-leading order.
- $f(p_i)$ means a generic parton/jet
- The processes include the decays of the W, Z and top quark, including spin correlations. Cuts can be applied to all particles.
- The particles are numbered by p_i, which is used in the output of the program to label the histograms

`input.DAT` : *basic set-up*

```
 '4.2'          [file version number]

[Flags to specify the mode in which MCFM is run]
 .false.        [evtgen]
 .false.        [creatent]
 .false.        [skipnt]
 .false.        [dswhisto]

[General options to specify the process and execution]
1               [nproc]
'tota'          [part 'lord','real' or 'virt','tota']
 'demo'         ['runstring']
14000d0         [sqrts in GeV]
+1              [ih1 =1 for proton and -1 for antiproton]
+1              [ih2 =1 for proton and -1 for antiproton]
 120d0          [hmass]
80d0            [scale:QCD scale choice]
80d0            [facscale:QCD fac_scale choice]
.false.         [dynamicscale]
 .false.        [zerowidth]
 .false.        [removebr]
 10             [itmx1, number of iterations for pre-conditioning]
 50000          [ncall1]
 10             [itmx2, number of iterations for final run]
 50000          [ncall2]
 1089           [ij]
 .false.        [dryrun]
 .false.        [Qflag]
 .true.         [Gflag]

[Pdf selection]
 'cteq6_m'      [pdlabel]
 4              [NGROUP, see PDFLIB]
 46             [NSET - see PDFLIB]
 cteq61.LHgrid  [LHAPDF group]
 -1             [LHAPDF set]
```

- Specify the process to study with the integer `nproc`.
- Set the order of the calculation: lowest order (`lord`), next-to-leading order real diagrams (`real`), virtual diagrams (`virt`) or the total NLO result (`tota`).
- Choose the collider energy (`sqrts`) and beam composition (`ih1,ih2`).
- Enter the values of the renormalization and factorization scales (`scale` and `facscale`).
- For some processes, choosing scales on an event-by-event basis is also possible (`dynamicscale`).

input.DAT : *integration parameters*

```
 '4.2'          [file version number]

[Flags to specify the mode in which MCFM is run]
 .false.        [evtgen]
 .false.        [creatent]
 .false.        [skipnt]
 .false.        [dswhisto]

[General options to specify the process and execution]
 1              [nproc]
 'tota'         [part 'lord','real' or 'virt','tota']
 'demo'         ['runstring']
 14000d0        [sqrts in GeV]
 +1             [ih1 =1 for proton and -1 for antiproton]
 +1             [ih2 =1 for proton and -1 for antiproton]
 120d0          [hmass]
 80d0           [scale:QCD scale choice]
 80d0           [facscale:QCD fac_scale choice]
 .false.        [dynamicscale]
.false.         [zerowidth]
 .false.        [removebr]
10              [itmx1, number of iterations for pre-conditioning]
50000           [ncall1]
10              [itmx2, number of iterations for final run]
50000           [ncall2]
1089            [ij]
 .false.        [dryrun]
 .false.        [Qflag]
 .true.         [Gflag]

[Pdf selection]
 'cteq6_m'      [pdlabel]
 4              [NGROUP, see PDFLIB]
 46             [NSET - see PDFLIB]
 cteq61.LHgrid  [LHAPDF group]
 -1             [LHAPDF set]
```

- Choose the description of resonances in the calculation (`zerowidth`): masses generated exactly at the peak (`true`) or off-shell (`false`). Important for $Z/\gamma^{\star}$.
- Choose the number of integration sweeps (`itmx1`) and points (`ncall1`) used in the warm-up stage of the calculation. `ncall1` should be increased when studying processes involving more particles.
- `itmx2` and `ncall2` are the values used when calculating all results.
- To produce distinct runs with otherwise-identical inputs, modify the random number seed `ij`.

`input.DAT` : *PDF specification*

```
 '4.2'           [file version number]

[Flags to specify the mode in which MCFM is run]
 .false.         [evtgen]
 .false.         [creatent]
 .false.         [skipnt]
 .false.         [dswhisto]

[General options to specify the process and execution]
 1               [nproc]
 'tota'          [part 'lord','real' or 'virt','tota']
 'demo'          ['runstring']
 14000d0         [sqrts in GeV]
 +1              [ih1 =1 for proton and -1 for antiproton]
 +1              [ih2 =1 for proton and -1 for antiproton]
 120d0           [hmass]
 80d0            [scale:QCD scale choice]
 80d0            [facscale:QCD fac_scale choice]
 .false.         [dynamicscale]
 .false.         [zerowidth]
 .false.         [removebr]
 10              [itmx1, number of iterations for pre-conditioning]
 50000           [ncall1]
 10              [itmx2, number of iterations for final run]
 50000           [ncall2]
 1089            [ij]
 .false.         [dryrun]
 .false.         [Qflag]
 .true.          [Gflag]

[Pdf selection]
'cteq6_m'        [pdlabel]
4                [NGROUP, see PDFLIB]
46               [NSET - see PDFLIB]
cteq61.LHgrid    [LHAPDF group]
-1               [LHAPDF set]
```

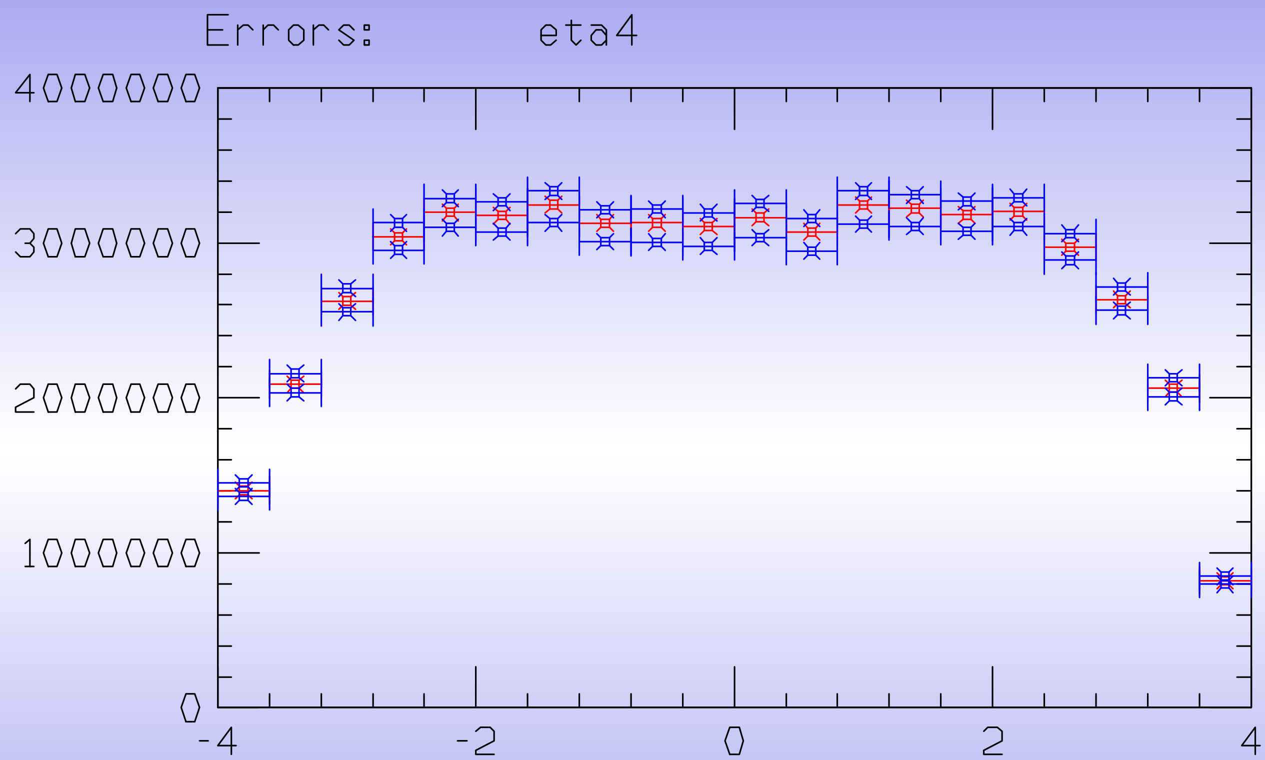
- The PDF set can be specified in one of 3 ways, depending on how the program is compiled
- The built-in implementations of popular MRS and CTEQ sets can be accessed via `pdlabel`.
- Access to PDFLIB is via the variables `NGROUP` and `NSET`.
- LHAPDF (successor to PDFLIB, see http://hepforge.cedar.ac.uk/lhapdf/) is the final option. Using this, the program can be used to evaluate PDF uncertainties (`LHAPDF set=-1`).

PDF demonstration

```
************ PDF error analysis ************
*                                          *
* PDF error set  0 ---> 11863519.187 fb    *
* PDF error set  1 ---> 11736647.263 fb    *
* PDF error set  2 ---> 11987706.935 fb    *
* PDF error set  3 ---> 11642296.578 fb    *
* PDF error set  4 ---> 12106702.170 fb    *
* PDF error set  5 ---> 11874355.602 fb    *
* PDF error set  6 ---> 11849124.661 fb    *
* PDF error set  7 ---> 11854123.307 fb    *
* PDF error set  8 ---> 11870788.530 fb    *
* PDF error set  9 ---> 11573575.545 fb    *
* PDF error set 10 ---> 12215308.143 fb    *
* PDF error set 11 ---> 12048942.634 fb    *
* PDF error set 12 ---> 11691755.691 fb    *
* PDF error set 13 ---> 11874942.354 fb    *
* PDF error set 14 ---> 11853027.865 fb    *
* PDF error set 15 ---> 11758576.633 fb    *
* PDF error set 16 ---> 11913098.358 fb    *
* PDF error set 17 ---> 11779363.844 fb    *
* PDF error set 18 ---> 11901319.122 fb    *
* PDF error set 19 ---> 12010652.500 fb    *
* PDF error set 20 ---> 11740933.791 fb    *
* PDF error set 21 ---> 11985450.352 fb    *
* PDF error set 22 ---> 11890845.603 fb    *
* PDF error set 23 ---> 11920163.560 fb    *
* PDF error set 24 ---> 11957587.988 fb    *
* PDF error set 25 ---> 11866956.994 fb    *
* PDF error set 26 ---> 11984564.841 fb    *
* PDF error set 27 ---> 11750019.234 fb    *
* PDF error set 28 ---> 11721240.685 fb    *
* PDF error set 29 ---> 11987664.142 fb    *
* PDF error set 30 ---> 11484760.388 fb    *
* PDF error set 31 ---> 11659901.056 fb    *
* PDF error set 32 ---> 11957864.136 fb    *
* PDF error set 33 ---> 11962671.598 fb    *
* PDF error set 34 ---> 11964869.177 fb    *
* PDF error set 35 ---> 11748873.948 fb    *
* PDF error set 36 ---> 11704653.918 fb    *
* PDF error set 37 ---> 11693637.274 fb    *
* PDF error set 38 ---> 11745558.021 fb    *
* PDF error set 39 ---> 11803672.011 fb    *
* PDF error set 40 ---> 11795692.786 fb    *
*                                          *
* --------------- SUMMARY ---------------- *
*                                          *
*      Minimum value  11484760.388 fb      *
*      Central value  11863519.187 fb      *
*      Maximum value  12215308.143 fb      *
*                                          *
*   Err estimate +/-    570618.935 fb      *
*      +ve direction    587609.633 fb      *
*      -ve direction    720774.743 fb      *
********************************************
```

- Sample output from a PDF uncertainty run. The program has used the CTEQ6 uncertainty sets (1 central + 40 others) and calculates W^+ production at NLO at the LHC.
- Rather than running the program 41 separate times, the cross section is evaluated for all PDF uncertainty sets at once. This is typically about 3–4 times slower than a normal run.
- The cross section is reported for each PDF set, then summarized at the end. The uncertainties are calculated according to the method laid out in `hep-ph/0101032`.
- In this case the uncertainty in the cross section due to the PDF is about 5%. This is a fairly typical result.

PDF uncertainties



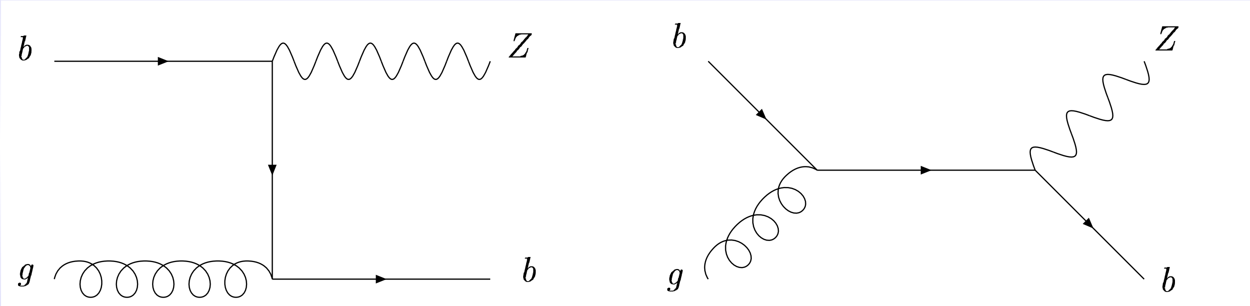
- Lepton pseudo-rapidity distribution at NLO.
- Central values are in red, extremal PDF values in blue. Note that in each bin, the PDF uncertainty set providing the extreme value may be different.

Application 1:
Single-tagged heavy flavour

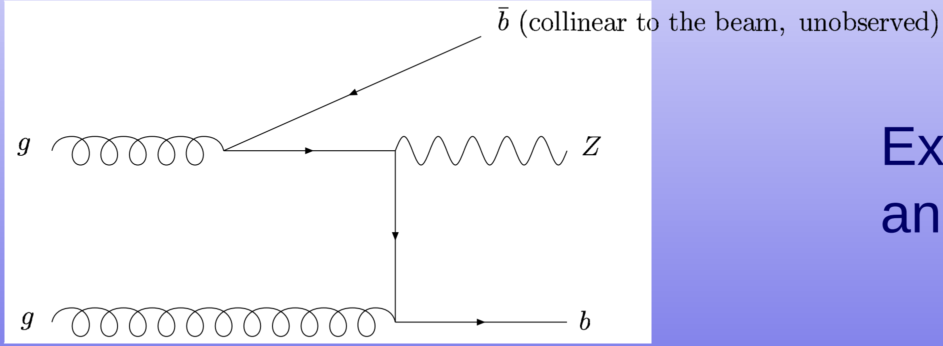
JC, Ellis, Maltoni, Willenbrock

Heavy flavour fraction revisited

- Often the presence of two b-quarks in the final state is actually only inferred from a single b-tag
- In this case, there is another way of computing the theoretical cross-section. For instance, in the case of $Z+$ heavy flavour:



- Requires knowledge of b-quark pdf's, but compare to:



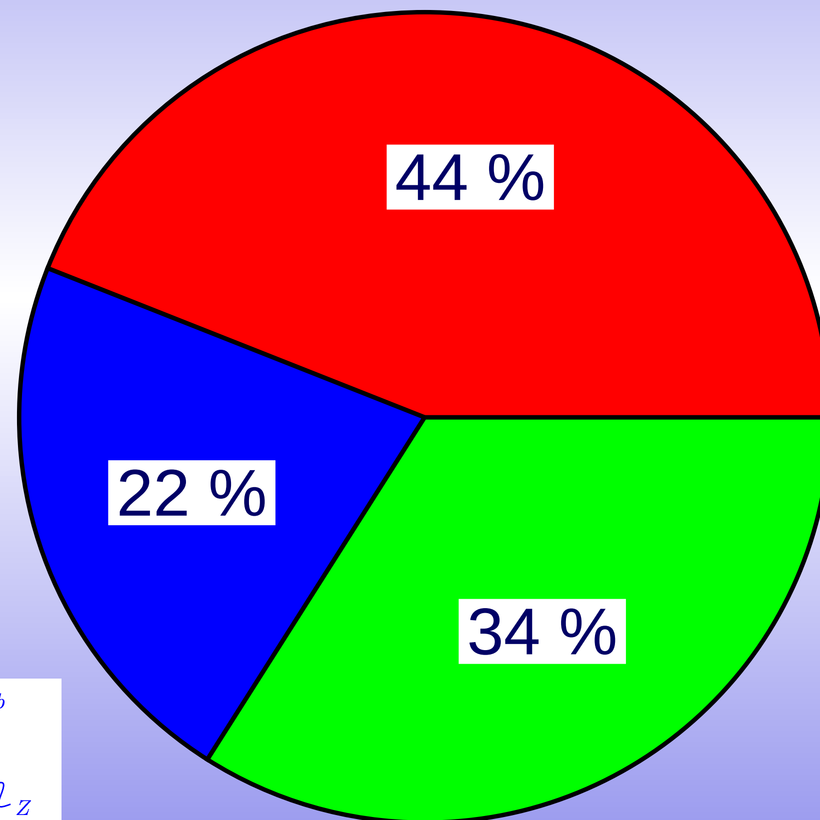
Expansion in $\alpha_s \ln(M_Z/m_b)$ and NLO calculation difficult

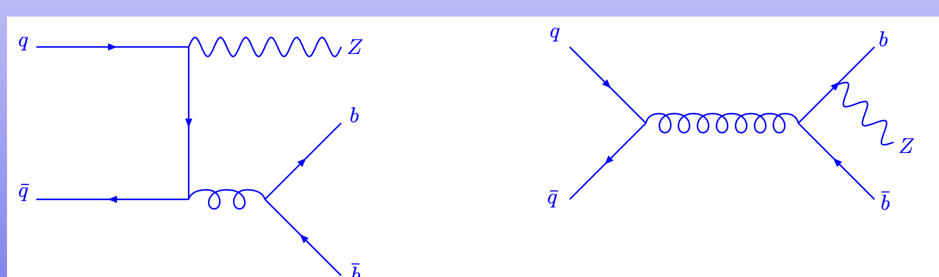
$Z + b$ *at NLO - Run II*

JC, K. Ellis, F. Maltoni and S. Willenbrock, hep-ph/0312024

- $p_T^{\rm jet} > 15$ GeV, $|\eta^{\rm jet}| < 2$
- $\sigma(Z + \text{ one } b \text{ tag}) = 20$ pb
- Fakes from $Z+$ jet events are significant
- Prediction for ratio of $Z + b$ to untagged $Z + \text{jet}$ is

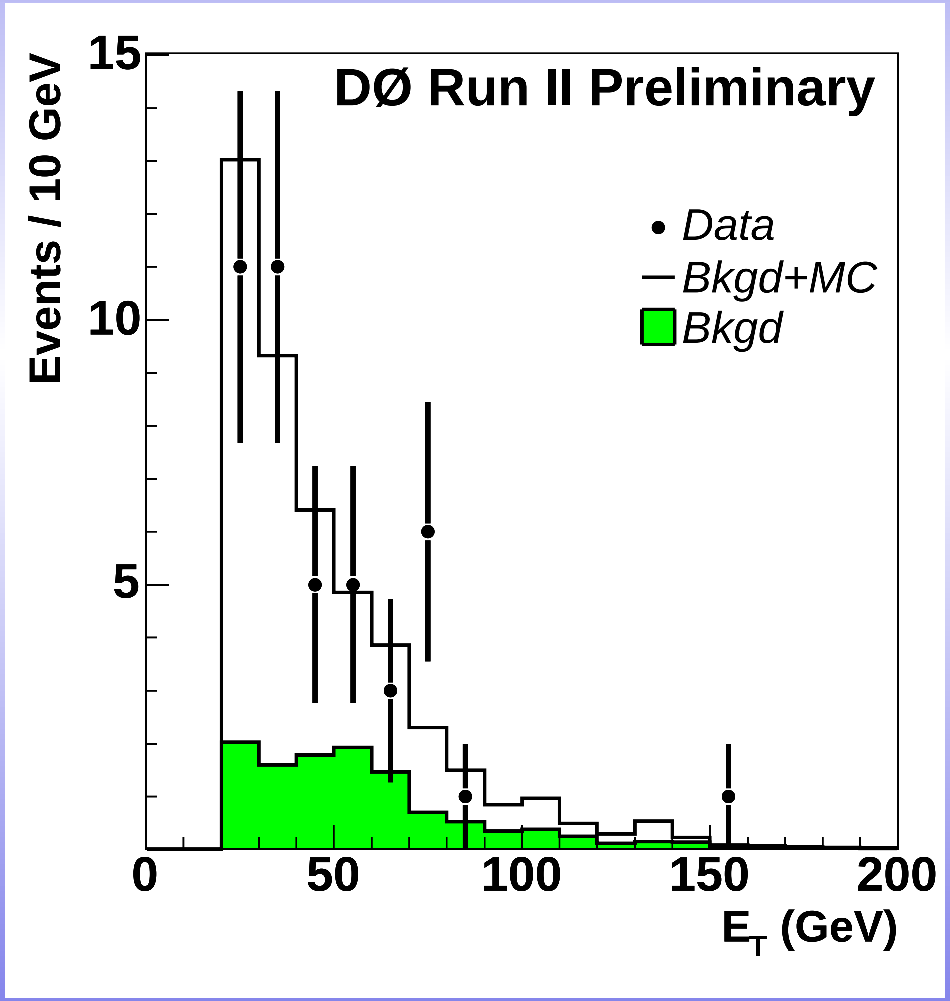
  0.02 ± 0.004





Experimental result

- Based on $189\ \mathrm{pb}^{-1}$ of data from Run II



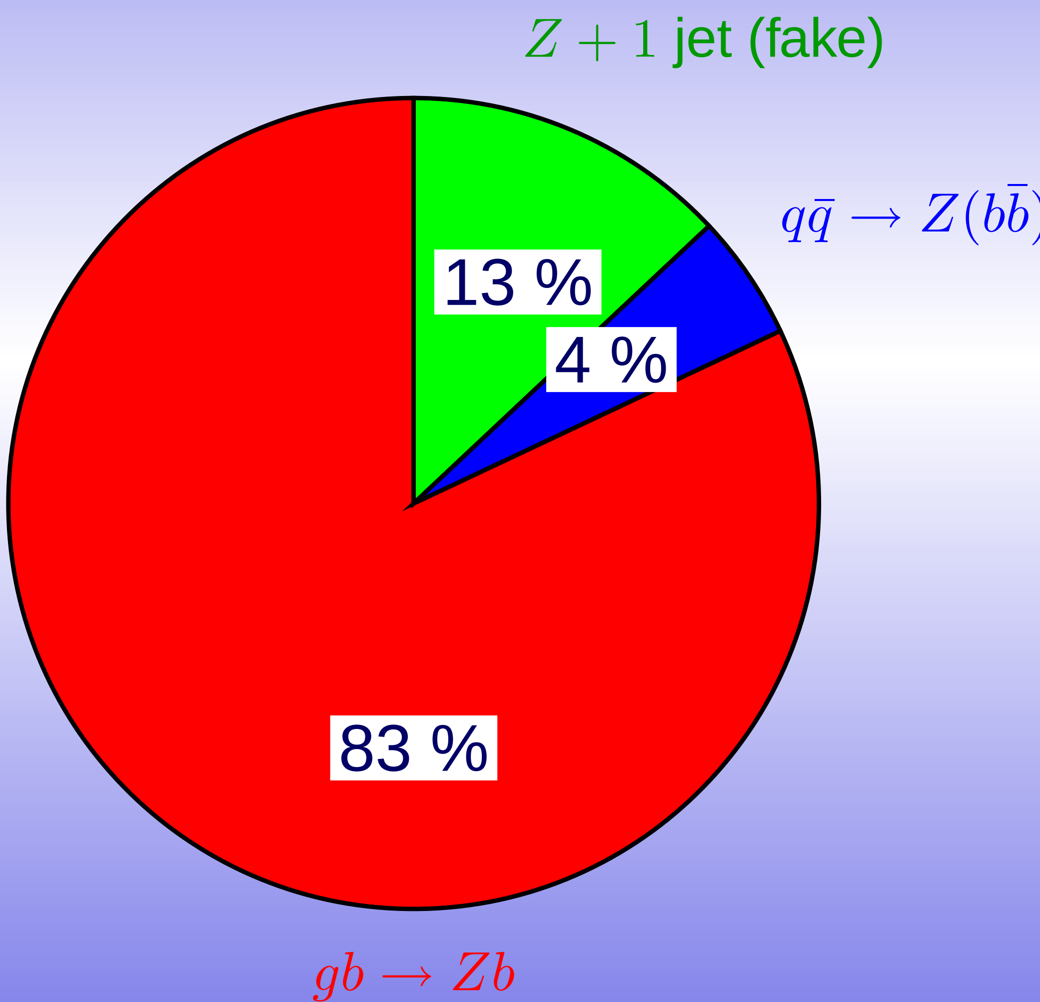
Ratio of cross-sections:

$$\frac{\sigma(Z+b)}{\sigma(Z+j)} = 0.024 \pm 0.007$$

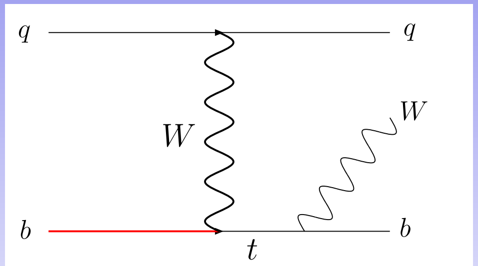
compatible with the NLO prediction from MCFM

LHC expectations

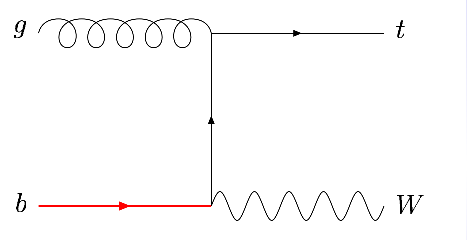
- $p_T^{\rm jet} > 15$ GeV, $|\eta^{\rm jet}| < 2.5$
- $\sigma(Z + \text{ one } b \text{ tag}) = 1$ nb
- Fakes from $Z+$ jet events are much less significant and $q\bar{q}$ contribution is tiny
- This should allow a fairly clean measurement of heavy quark PDF's (currently, only derived perturbatively)



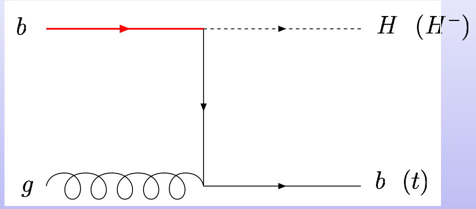
b-PDF uses



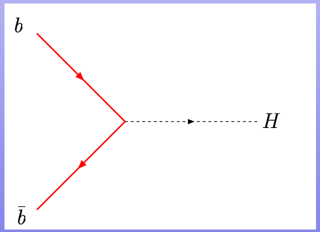
single-top $qb \to qWb$



single-top $gb \to tW$



(charged) Higgs+b



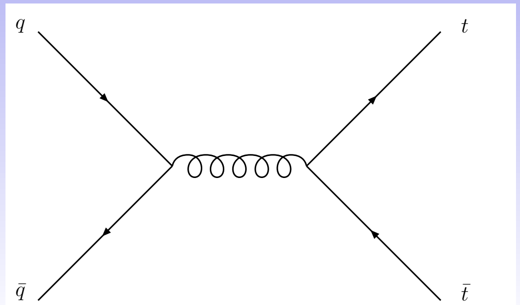
inclusive Higgs

Application 2:
Single top production and decay

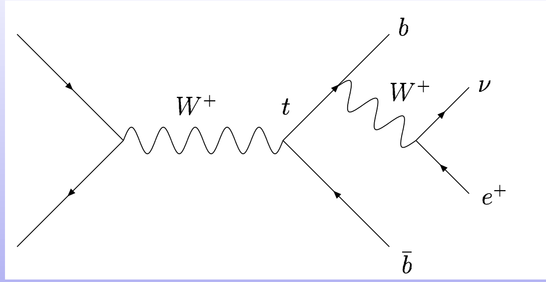
JC, Ellis, Tramontano

Producing the top quark

- The top quark was discovered in Run I of the Tevatron by producing it in pairs:

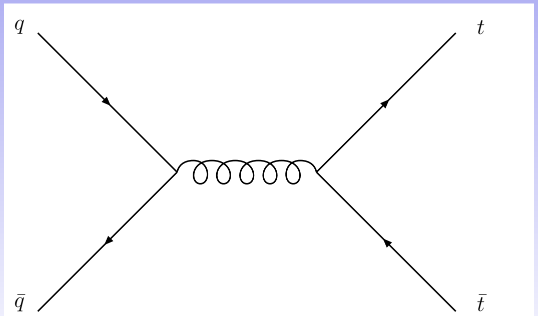


- However, it should also be possible to produce it singly in Run II, for example:

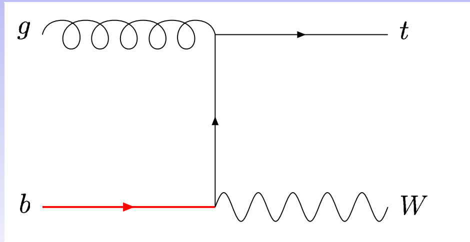


- This is especially interesting since it would yield information about the weak interaction of top quarks (V_{tb}).

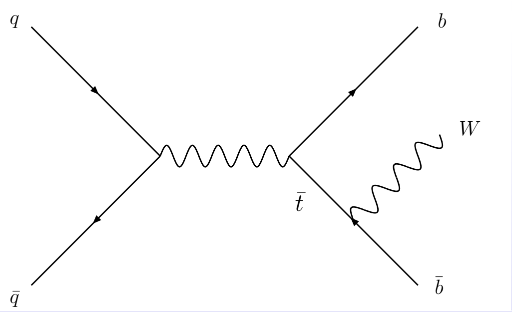
Top production rates



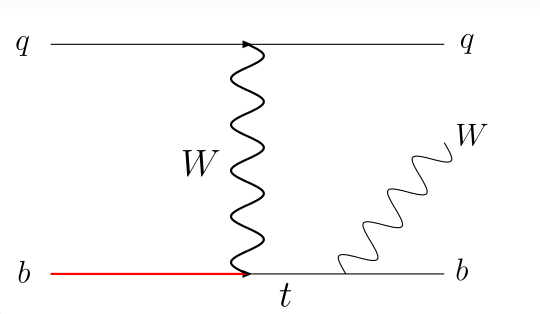
6 pb
720 pb



0.14 pb
66 pb



0.8 pb
10 pb

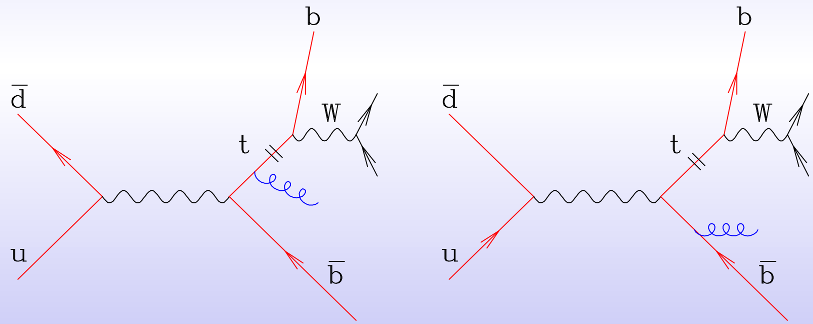


1.8 pb
240 pb

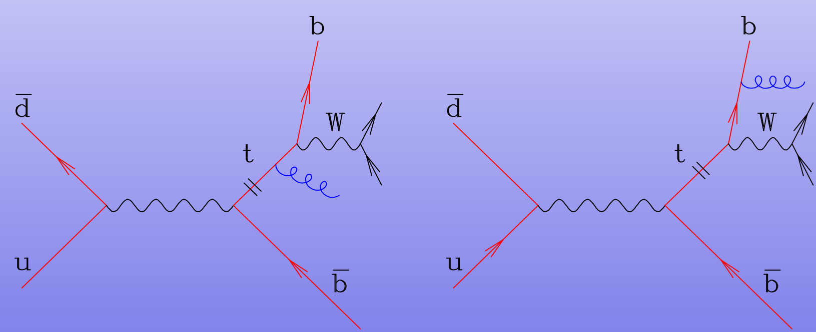
- All cross-sections are known to NLO (Tevatron / LHC)
- The total single top cross-section is smaller than the $t\bar{t}$ rate by about a factor of two, at both machines

Inclusion of decay

- Results had previously been presented without including the decay of the top quark. Without it, predictions for some quantities used in Tevatron search strategies are impossible
- Final state radiation that enters at next-to-leading order is possible in either the production or decay phase:

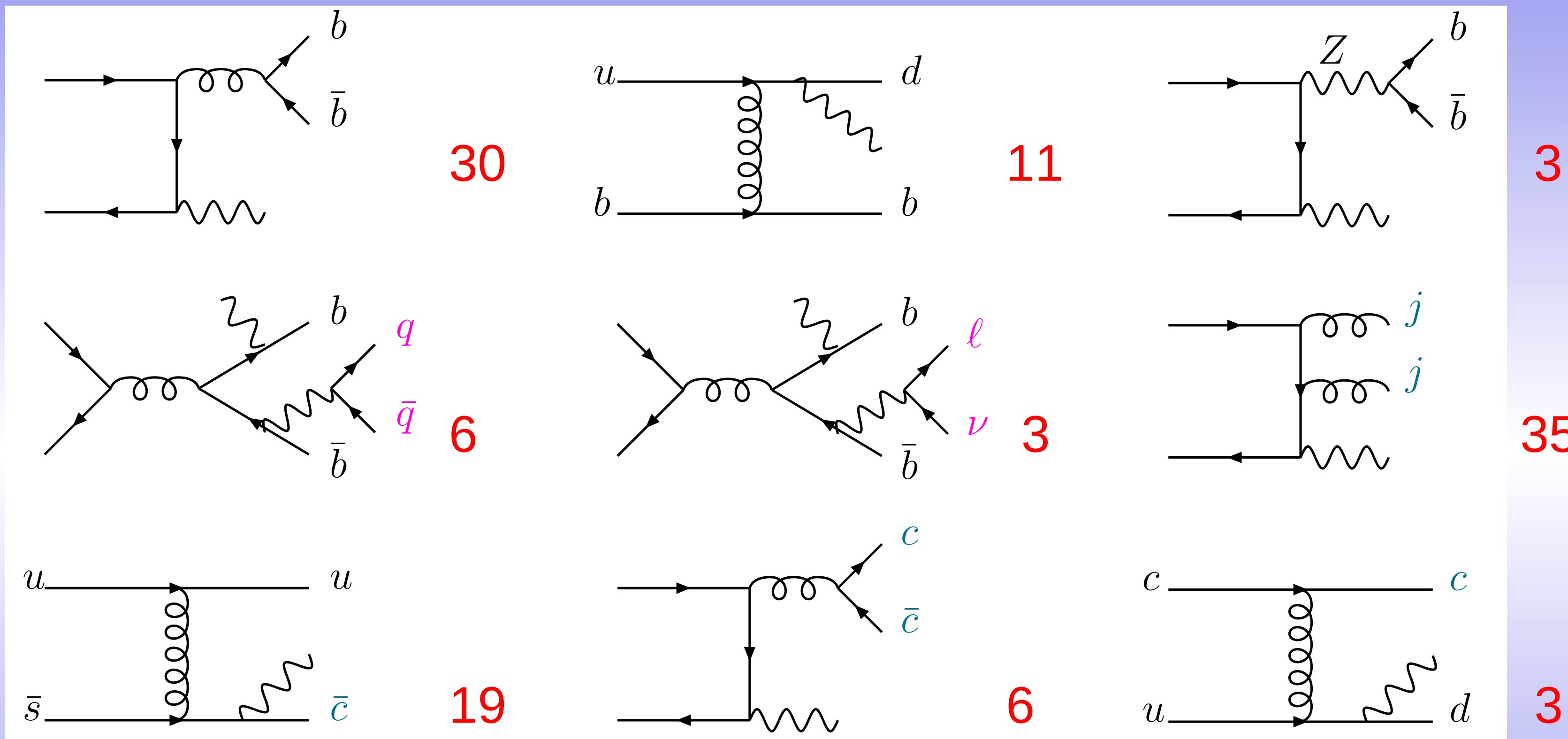


production



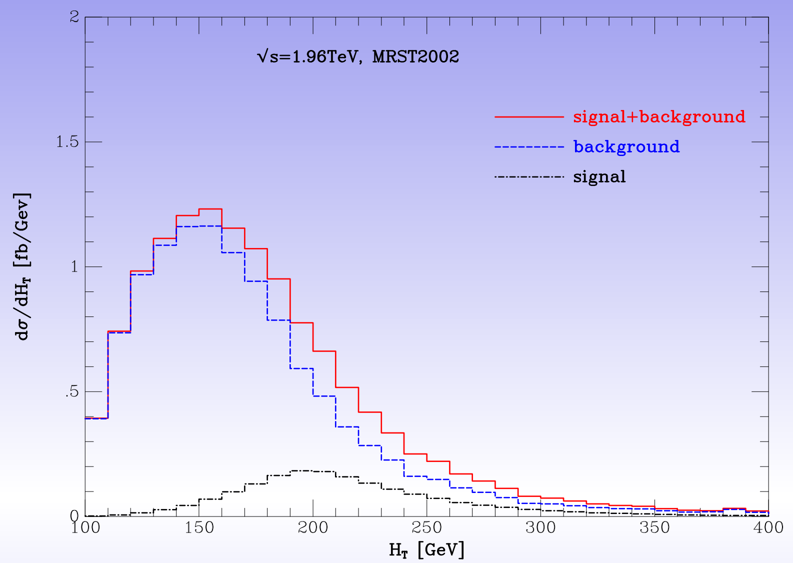
decay

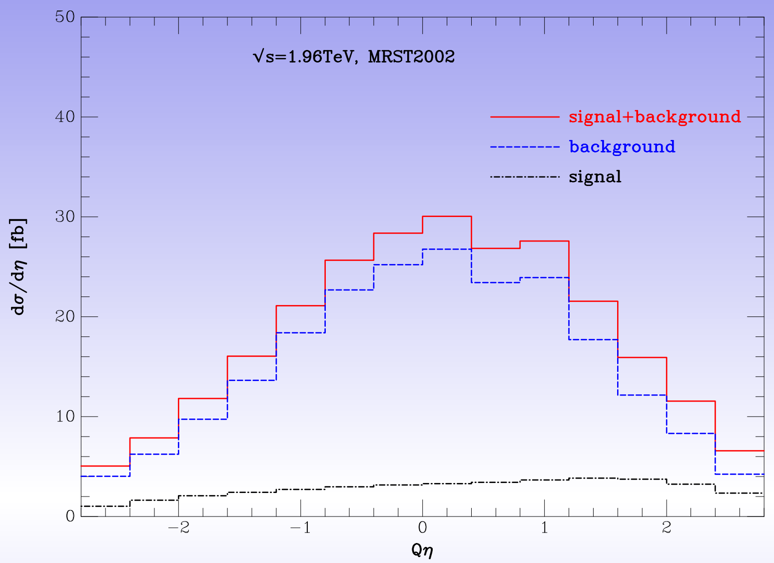
Backgrounds at the Tevatron



- Cross-sections in fb include nominal tagging efficiences and mis-tagging/fake rates. Calculated with MCFM, most at NLO
- Rates are 7 fb and 11 fb for s- and $t-$channel signal

Single top signal vs. backgrounds





- $H_T =$ scalar sum of jet, lepton and missing E_T
- $Q\eta$ is the product of the lepton charge and the rapidity of the untagged jet, useful for picking out the t-channel process
- Signal:Background (with our nominal efficiencies) is about $1:6$ – a very challenging measurement indeed. Production in this mode has not yet been observed at Fermilab.
- Knowing the characteristics of signal and background events at NLO should help. D0 estimate $7\ \mathrm{fb}^{-1}$ for a 5σ observation.

Shortcomings

The approach in MCFM involves a number of approximations:

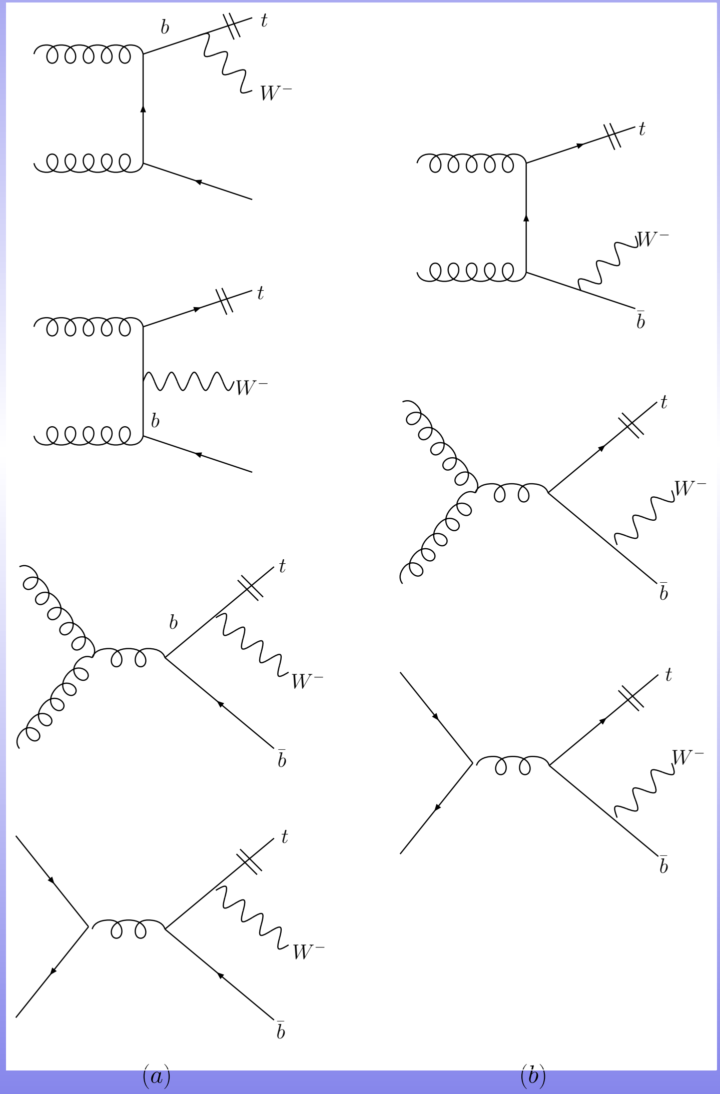
- The b-quark is massless
  LO calculation with $m_b = 4.75$ GeV $\longrightarrow < 1\%$ effect
- The top quark is put on its mass-shell
  LO calculation with a Breit-Wigner $\longrightarrow 1\%$ effect
- We neglect interference between radiation in production/decay
  qualitative argument for $\mathcal{O}(\alpha_s \Gamma_t / m_t) \sim$ less than a percent
- We assume p_T-independent heavy flavour tagging efficiencies, as well as stable b and c quarks
  easily addressed by a more detailed experimental analysis with the publicly-available code
- No showering or hadronization is performed
  no NLO/PS prediction yet available; however the large cone size $\Delta R = 1$ should help minimize these effects

Associated Wt production

JC, Tramontano

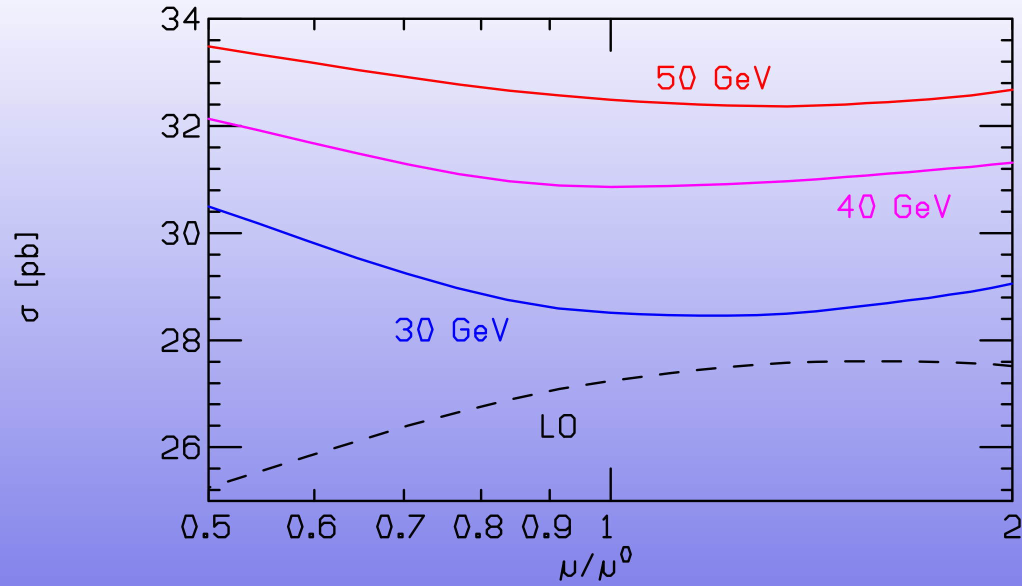
Details

- Calculation proceeds in the same way
- NLO real radiation corrections contain the process $gg \to Wtb$. This final state is also obtained by producing two top quarks on shell.
- Including this contribution doesn't give a meaningful result.
- Previous attempts to remove it involve either subtracting the resonant contribution or applying an invariant mass cut
- Neither of these is suitable for a MC approach including decays



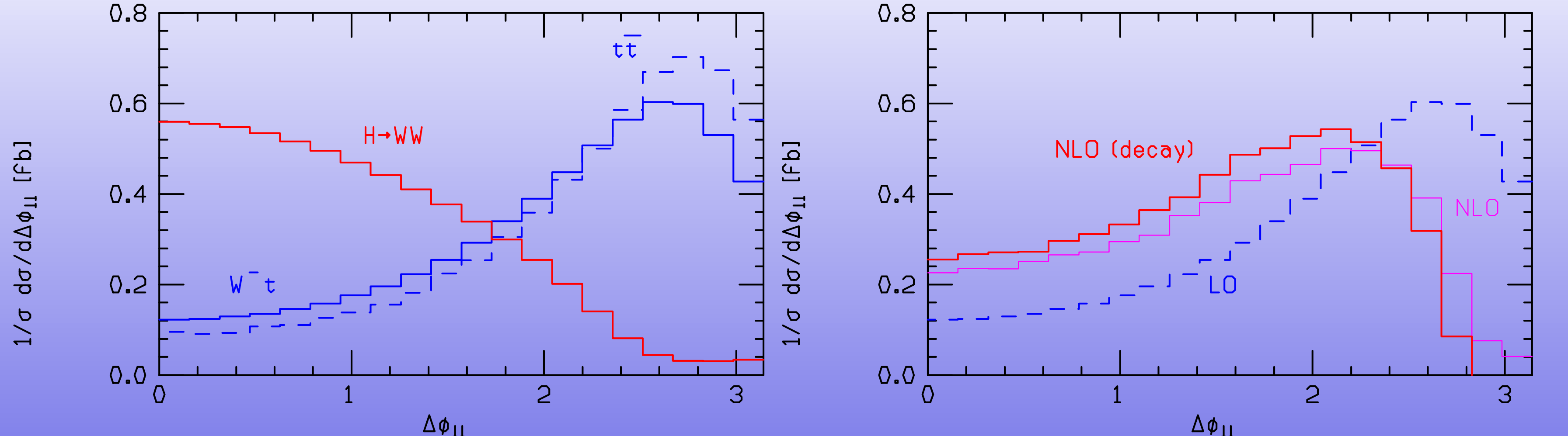
Solution

- Contribution when b quark p_T is small is accounted for by the b-quark PDF
- When the b quark p_T is large, the event should be best described by resonant $t\bar{t}$ production
- Therefore, define the Wt process by demanding that no b quark be observed above a given value of $p_T = \mu_V$. Factorization and renormalization scales should also be chosen equal to μ_V.
- As a result, the NLO prediction depends on the value chosen for μ_V
- Corrections are mild



LHC analysis: $gg \to H \to WW^*$

- $155 < m_H < 180$ GeV, so W's decay to leptons. Signal is two leptons and missing E_T
- Main background is from continuum W pair production, via $q\bar{q}$ scattering and loop-induced gluon-gluon fusion
- Further backgrounds from events containing leptonically-decaying top quarks where the jets are not observed
- Enhance signal using strong cuts. The opening angle between the leptons in the transverse plane is a good discriminator.



Summary

- MCFM is a general purpose NLO code, designed to be downloaded and run by users:

  `http://mcfm.fnal.gov/`

- Many of the included processes are backgrounds to new physics searches. Notable examples are diboson production, W/Z plus up to 2 jets and single top production.
- In general NLO provides a stable normalization of cross sections. The uncertainty can be investigated using MCFM, for example by varying input scales and using PDF error sets.
- Simple cuts are easy to apply; more complicated ones can be included with a little effort. By matching the cuts with the experimental ones, the effect of NLO corrections can be better assessed.