Lecture 3

QCD at the LHC
Joey Huston
Michigan State University
Why stop at NLO?

- NNLO is even better, but also more complicated
- Expect (even further) reduced scale dependence
- And in some cases, extra cross section contributions, such as for Higgs production
- Have only been carried out for a few processes to date
- Would really like to have it for inclusive jet production, for example
  - currently > 8 year project
Higgs: LO→NLO→NNLO


convergence in going:
LO → NLO → NNLO

Confirmed by the full scale dependence:

Note that the NLO result is not covered by the LO error band
All-orders approaches

- Rather than systematically calculating to higher and higher orders in the perturbative expansion, can also use a number of all-orders approaches
- In resummation, dominant contributions from each order in perturbation theory are singled out and resummed by use of an evolution equation
- Near boundaries of phase space, fixed order calculations break down due to large logarithmic corrections, and these contributions can become important. Resummation takes them into account.

Consider W production
- one large logarithm associated with production of vector boson close to threshold
  - takes form of \( \alpha_s^n \log^{2n-1}(1-z)/(1-z) \) where \( z = Q^2/s-1 \)
  - other large logarithm is associated with recoil of vector boson at very small \( p_T \)
    - logarithms appear as \( \alpha_s^n \log^{2n-1}(Q^2/p_T^2) \)

In both cases there is a restriction of phase space for gluon emission and thus the logs become large and are crucial for an accurate prediction.
All-orders approaches

• Remember the expression we had after adding gluons on to the $W + 1$ jet process
  - each gluon added yields an additional factor of $\alpha_s$ and two new logarithms

\[
d\sigma = \sigma_0(W + 1 \text{ jet}) \left[1 + \alpha_s(c_{12} L^2 + c_{11} L + c_{10}) + \alpha_s^2(c_{24} L^4 + c_{23} L^3 + c_{22} L^2 + c_{21} L + c_{20}) + \cdots \right]
\]

• $q_T$ resummation is resumming the effects of logs of $Q^2/p_T^2$

• note that $q_T$ resummation does not change the size of the cross section; it just modifies the $p_T$ distribution of the $W$

• These are the leading logs (LL) (highest power of log for each power of $\alpha_s$)

• These are the next-to-leading logs (NLL) (next highest power of log…)
  - …and so on

• We know the structure of the LL’s, NLL’s, NNLL’s

• But we don’t know the $c_{ij}$ factors until we do the finite order calculation

• LO gives us the LL

• NLO gives us the NLL
  - …and so on

• The accuracy of the resummation improves with the addition for further higher order information

• A resummation program like ResBos has NNLL accuracy
All-orders approaches

- Remember the expression we had after adding gluons on to the $W + 1$ jet process
  - each gluon added yields an additional factor of $\alpha_s$ and two new logarithms

\[
\frac{d\sigma}{dp_T^2} = \sigma \frac{d}{dp_T^2} \exp \left( -\frac{\alpha_s C_F}{2\pi} \log^2 \frac{M_W^2}{p_T^2} \right)
\]

- $q_T$ resummation is resumming the effects of logs of $Q^2/p_T^2$

- note that $q_T$ resummation does not change the size of the cross section; it just modifies the $p_T$ distribution of the $W$

Note that distribution goes to zero as $p_T \to 0$; no divergence

Figure 20. The resummed (leading log) $W$ boson transverse momentum distribution.

You could get the same predictions by using PDFs in which the transverse momentum ($k_T$) has not been integrated out
Parton showers

- A different, but related approach for re-summing logarithms, is provided by parton showering.
- By the use of the parton showering process, a few partons produced in a hard interaction at a high-energy scale can be related to partons at an energy scale close to $\Lambda_{\text{QCD}}$.
- At this lower energy scale, a universal non-perturbative model can then be used to provide the transition to hadrons.
- Parton showering allows for evolution, using DGLAP formalism, of parton fragmentation function.

Parton Cascade

...plus similar for initial state

- Due to successive branching, parton cascade or shower develops. Each outgoing line is source of new cascade, until all outgoing lines have stopped branching. At this stage, which depends on cutoff scale $t_0$, outgoing partons have to be converted into hadrons via a hadronization model.
- Successive values of an evolution variable $t$, a momentum fraction $z$, and an azimuthal angle $\phi$ are generated, along with the flavors of the partons emitted during the parton shower.
Parton shower evolution

- On average, emitted gluons have decreasing angles with respect to parent parton directions
  - angular ordering, an aspect of color coherence
- The evolution variable $t$ can be the virtuality of the parent parton [old Pythia and old Sherpa], $E^2(1-\cos\theta)$ where $E$ is the energy of the parent parton and $\theta$ is the opening angle between the two partons [Herwig], or the square of the transverse momentum between the two partons [new Pythia]
Sudakov form factors

- Sudakov form factors form the basis for both resummation and parton showering.
- We can write an expression for the Sudakov form factor of an initial state parton in the form below, where $t$ is the hard scale, $t_0$ is the cutoff scale and $P(z)$ is the splitting function.

$$\Delta(t) \equiv \exp \left[ - \int_{t_0}^{t} \frac{dt'}{t'} \int \frac{dz}{z} \frac{\alpha_S}{2\pi} P(z) \frac{f(x/z, t)}{f(x, t)} \right]$$

- Similar form for the final state but without the pdf weighting.
- Sudakov form factor resums all effects of soft and collinear gluon emission (so again the double logs), but does not include non-singular regions that are due to large energy, wide angle gluon emission.
- Gives the probability **not** to radiate a gluon greater than some energy.
- We can draw explicit (approximate) curves for the Sudakov form factors.

Figure 21. The Sudakov form factors for initial-state gluons at a hard scale of 100 GeV as a function of the transverse momentum of the emitted gluon. The form factors are for (top to bottom) parton $x$ values of 0.3, 0.1, 0.03, 0.01, 0.001 and 0.0001.

Figure 22. The Sudakov form factors for initial-state gluons at a hard scale of 500 GeV as a function of the transverse momentum of the emitted gluon. The form factors are for (top to bottom) parton $x$ values of 0.3, 0.1, 0.03, 0.01, 0.001 and 0.0001.
The Sudakov form factor decreases (the probability of radiating increases) as the $p_T$ of the radiated gluon decreases, as the hardness of the interaction increases, or as the $x$ value of the incoming parton decreases (more phase space for gluon radiation).

**Figure 21.** The Sudakov form factors for initial-state gluons at a hard scale of 100 GeV as a function of the transverse momentum of the emitted gluon. The form factors are for (top to bottom) parton $x$ values of 0.3, 0.1, 0.03, 0.01, 0.001 and 0.0001.

**Figure 22.** The Sudakov form factors for initial-state gluons at a hard scale of 500 GeV as a function of the transverse momentum of the emitted gluon. The form factors are for (top to bottom) parton $x$ values of 0.3, 0.1, 0.03, 0.01, 0.001 and 0.0001.
Sudakov form factors: quarks and gluons

so quarks don’t radiate as much; it’s the $C_F$ compared to $C_A$

Figure 23. The Sudakov form factors for initial-state quarks at a hard scale of 100 GeV as a function of the transverse momentum of the emitted gluon. The form factors are for (top to bottom) parton $x$ values of 0.3, 0.1 and 0.03.

Figure 21. The Sudakov form factors for initial-state gluons at a hard scale of 100 GeV as a function of the transverse momentum of the emitted gluon. The form factors are for (top to bottom) parton $x$ values of 0.3, 0.1, 0.03, 0.01, 0.001 and 0.0001.

Figure 24. The Sudakov form factors for initial-state quarks at a hard scale of 500 GeV as a function of the transverse momentum of the emitted gluon. The form factors are for (top to bottom) parton $x$ values of 0.3, 0.1 and 0.03.

Figure 22. The Sudakov form factors for initial-state gluons at a hard scale of 500 GeV as a function of the transverse momentum of the emitted gluon. The form factors are for (top to bottom) parton $x$ values of 0.3, 0.1, 0.03, 0.01, 0.001 and 0.0001.
Let’s take a specific example

- Consider gg fusion production of a 120 GeV Higgs (so use Q=100 GeV) at 7 TeV.
- The x values of each of the incoming gluons are in the range from 0.01 to 0.03.
- For an ISR jet (gluon) with a $p_T$ of 20 GeV/c or greater, the Sudakov form factor is on the order of 0.7, i.e. there is a 30% chance of radiating such a gluon.
- Since there are two incoming gluons, chance of having a jet > 20 GeV/c is \(~0.5\)

Figure 21. The Sudakov form factors for initial-state gluons at a hard scale of 100 GeV as a function of the transverse momentum of the emitted gluon. The form factors are for (top to bottom) parton $x$ values of 0.3, 0.1, 0.03, 0.01, 0.001 and 0.0001.
We can only observe emissions above a certain resolution scale.

Below this resolution scale, singularities cancel, leaving a finite remnant.

(some of) the virtual corrections encountered in a full NLO calculation are included by the use of Sudakov suppression between vertices.

So a parton shower Monte Carlo is not purely a fixed order calculation, but has a higher order component as well.

This is a statement that you’ll often hear.
Merging ME and PS approaches

- Parton showers provide an excellent description in regions which are dominated by soft and collinear gluon emission
- Matrix element calculations provide a good description of processes where the partons are energetic and widely separated and also take into account interference effects between amplitudes
  - but do not take into account interference effects in soft and collinear emissions which cannot be resolved, and thus lead to Sudakov suppression of such emissions
- Hey, I know, let’s put them together, but we have to be careful not to double-count
  - parton shower producing same event configuration already described by matrix element
  - Les Houches Accord (which I named) allows the ME program to talk to the PS program

Figure 26. In the NLO formalism, the same scale, proportional to the hardness of the process, is used for each QCD vertex. For the case of the $W+2$ jet diagram shown above to the left, a scale related to the mass of the $W$ boson, or to the average transverse momentum of the produced jets, is typically used. The figure to the right shows the results of a simulation using the CKKW formalism. Branchings occur at the vertices with resolution parameters $d_i$, where $d_1 > d_2 > d_3 > d_4 > d_5 > d_6$. Branchings at the vertices 1–2 are produced with matrix element information while the branchings at vertices 3–6 are produced by the parton shower.
Merging ME and PS approaches

- A number of techniques to combine, with most popular/correct being CKKW
  - Matrix element description used to describe parton branchings at large angle and/or energy
  - Parton shower description is used for smaller angle, lower energy emissions

- Division into two regions of phase space provided by a resolution parameter $d_{\text{ini}}$

- Argument of $\alpha_s$ at all of the vertices is taken to be equal to the resolution parameter $d_i$ (showering variable) at which the branching has taken place

- Sudakov form factors are inserted on all of the quark and gluon lines to represent the lack of any emissions with a scale larger than $d_{\text{ini}}$ between vertices
  - Parton showering is used to produce additional emissions at scales less than $d_i$

- For typical matching scale, $\sim 10\%$ of the n-jet cross section is produced by parton showering from n-1 parton ME

---

Figure 26. In the NLO formalism, the same scale, proportional to the hardness of the process, is used for each QCD vertex. For the case of the $W+2$ jet diagram shown above to the left, a scale related to the mass of the $W$ boson, or to the average transverse momentum of the produced jets, is typically used. The figure to the right shows the results of a simulation using the CKKW formalism. Branchings occur at the vertices with resolution parameters $d_i$, where $d_1 > d_2 > d_3 > d_4 > d_5 > d_6$. Branchings at the vertices 1–2 are produced with matrix element information while the branchings at vertices 3–6 are produced by the parton shower.

See Alpgen, Madgraph, Sherpa,... MLM approach (which I also named) is an approximation to the full CKKW procedure
Best of all worlds: from CHS

NLO + parton shower MC

Several groups have worked on the subject to consistently combine partonic NLO calculations with parton showers.

- Collins, Zu [95, 96].
- Frixione, Nason, Webber (MC@NLO) [97–99].
- Kurihara, Fujimoto, Ishikawa, Kato, Kawabata, Munehsia, Tanaka [100].
- Krämer, Soper [101–103].
- Nagy, Soper [104, 105].

MC@NLO is the only publicly available program that combines NLO calculations with parton showering and hadronization. The HERWIG Monte Carlo is used for the latter. The use of a different Monte Carlo, such as PYTHIA, would require a different subtraction scheme for the NLO matrix elements. The processes included to date are \((W, Z, \gamma^*, H, bb, t\bar{t}, HW, HZ, WW, WZ, ZZ)\). Recently, single top hadroproduction has been added to MC@NLO [106]. This is the first implementation of a process that has both initial- and final-state singularities. This allows a more general category of additional processes to be added in the future. Work is proceeding on the addition of inclusive jet production and the production of a Higgs boson via \(WW\) fusion. Adding spin correlations to a process increases the level of difficulty but is important for processes such as single top production.

If, in addition, the CKKW formalism could be used for the description of hard parton emissions, the utility and accuracy of a NLO Monte Carlo could be greatly increased. The merger of these two techniques should be possible in Monte Carlos available by the time of the LHC turn-on.

Powheg is a program/technique that results in no negative weight events (or a few); many new processes have been added

But it’s still very time(theorist)-consuming to add a new process, and most of the processes so far are 2->1, 2->2
At low $p_T$ for the t-tbar system, the cross section is described (correctly) by the parton shower, which resums the large logs near $p_T \sim 0$.

At high $p_T$, the cross section is described (correctly) by the NLO matrix element.
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.
Global fits

- With the DGLAP equations, we know how to evolve pdf’s from a starting scale $Q_0$ to any higher scale
  - remember the divergences from the initial state that we absorbed into the pdfs
- ...but we can’t calculate what the pdf’s are ab initio
  - one of the goals of lattice QCD
- We have to determine them from a global fit to data
  - factorization theorem tells us that pdf’s determined for one process are applicable to another

- So what do we need
  - a value of $Q_0$ (1.3 GeV for CTEQ, 1 GeV for MSTW) lower than the data used in the fit (or any prediction)
  - a parametrization for the pdf’s
  - a scheme for the pdf’s
  - hard-scattering calculations at the order being considered in the fit
  - pdf evolution at the order being considered in the fit
  - a world average value for $\alpha_s$
  - a lot of data
    - with appropriate kinematic cuts
  - a treatment of the errors for the experimental data
  - MINUIT
Back to global fits

- **Parametrization: initial form**
  - $f(x) \sim x^\alpha (1-x)^\beta$
  - estimate $\beta$ from quark counting rules
    - $\beta = 2n_s - 1$ with $n_s$ being the minimum number of spectator quarks
    - so for valence quarks in a proton (qqq), $n_s = 2$, $\beta = 3$
    - for gluon in a proton (qqqg), $n_s = 3$, $\beta = 5$
    - for anti-quarks in a proton (qqqqqbar), $n_s = 4$, $\beta = 7$
  - estimate $\alpha$ from Regge arguments
    - gluons and anti-quarks have $\alpha \sim -1$ while valence quarks have $\alpha \sim 1/2$
  - but at what $Q$ value are these arguments valid?

- **What do we know?**
  1. we know that the sum of the momentum of all partons in the proton is 1 (but not for modified LO fits)
  2. we know the sum of valence quarks is 3
    - and 2 of them are up quarks and 1 of them is a down quark
    - we know that the net number of anti-quarks is 0, but what about $d \bar{d} = u \bar{u}\bar{d}$
  3. we know that the net number of strange quarks (charm quarks/bottom quarks) in the proton is 0
    - but we don’t know if $s = s\bar{s}$ locally

This already puts a lot of restrictions on the pdf’s
Parametrizations

- That simple parametrization worked for early fits, where the data was not very precise (nor very abundant), but it does not work for modern global fits, where a more flexible form is needed
  - the simple ansatz can be dangerous in that it can (falsely) tie together low x and high x behavior (other than by momentum sum rule)
- In order to more finely tune parametrization, usually multiply simple form by a polynomial in x or some more complicated function

- CTEQ uses for the quark and gluon distributions (CTEQ6.6)

$$f(x) = x^{(a_1-1)}(1-x)^{a_2} e^{a_3 x} [1 + e^{a_4 x}]^{a_5}$$

- For the ratio of dbar/ubar

$$\frac{\bar{d}}{\bar{u}} = e^{a_1} x^{(a_2-1)}(1-x)^{a_3} + (1 + a_4 x)(1 - x)^{a_5}$$

- How do we know this is flexible enough?
  - data is well-described ($\chi^2$/dof ~1 for a NLO fit)
  - adding more parameters just results in those parameters being unconstrained
  - but there is some remaining bias
  - note that with this form, the pdf’s are positive definite (they don’t have to be)
Orders and Schemes

- Fits are available at
  - LO
    ▲ CTEQ6L or CTEQ6L1
      – 1 loop or 2 loop $\alpha_s$
    ▲ in common use with parton shower Monte Carlos
  ▲ poor fit to data due to deficiencies of LO ME’s
  - LO*
    ▲ better for parton shower Monte Carlos
  - NLO
    ▲ CTEQ6.1, CTEQ6.6, CT09
    ▲ precision level: error pdf’s defined at this order
  - NNLO
    ▲ more accurate but not all processes known

- At NLO and NNLO, one needs to specify a scheme or convention for subtracting the divergent terms

- Basically the scheme specifies how much of the finite corrections to subtract along with the divergent pieces
- most widely used is the modified minimal subtraction scheme (or MSbar)
- used with dimensional regularization: subtract the pole terms and accompanying $\log 4\pi$ and Euler constant terms

\[ \sigma_{\text{real+corr}} = \frac{\alpha_s}{2\pi} C_A \left( \frac{\mu^2}{Q^2} \right)^\gamma \left[ \left( \frac{2\pi^2}{3} - 6 \right) \delta(1-z) + \frac{2}{\varepsilon} P_{\text{corr}}(z) - 2(1-z) + 4(1+z) \left( \frac{\ln(1-z)}{1-z} \right) - 2 \right] \]

- also may find pdf’s in DIS scheme, where full order $\alpha_s$ correction for $F_2$ in DIS absorbed into quark pdf’s
Scales and Masses

- Processes used in global fits are characterized by a single large scale:
  - DIS-$Q^2$
  - lepton pair production-$M^2$
  - vector boson production-$M_{\nu}^2$
  - jet production-$p_{T}^{\text{jet}}$

- By choosing the factorization and renormalization scales to be of the same order as the characteristic scale:
  - can avoid some large logarithms in the hard scattering cross section
  - some large logarithms in running coupling and pdf’s are resummed

Different treatment of quark masses and thresholds:

- fixed flavor number scheme (FFNS)
- variable flavor number scheme (VFNS)
  - zero mass variable flavor number scheme (ZM-VFNS)
  - general mass variable flavor number scheme (GM-VFNS)
Data sets used in global fits (CTEQ6.6)

1. BCDMS $F_2^{proton}$ (339 data points)
2. BCDMS $F_2^{deuteron}$ (251 data points)
3. NMC $F_2$ (201 data points)
4. NMC $F_2^o/F_2^p$ (123 data points)
5. $F_2$ (CDHSW) (85 data points)
6. $F_2$ (CDHSW) (96 data points)
7. CCFR $F_2$ (69 data points)
8. CCFR $F_3$ (86 data points)
9. H1 NC e$^-$p (126 data points; 1998-98 reduced cross section)
10. H1 NC e$^-$p (13 data points; high $y$ analysis)
11. H1 NC e$^+$p (115 data points; reduced cross section 1996-97)
12. H1 NC e$^+$p (147 data points; reduced cross section; 1999-00)
13. ZEUS NC e$^+$p (92 data points; 1998-99)
14. ZEUS NC e$^+$p (227 data points; 1996-97)
15. ZEUS NC e$^+$p (90 data points; 1999-00)
16. H1 $F_2^c$ e$^+$p (8 data points; 1996-97)
17. H1 R$^c$ for ccbar e$^+$p (10 data points; 1996-97)
18. H1 R$^b$ for bbbar e$^+$p (10 data points; 1999-00)
19. ZEUS $F_2^c$ e$^+$p (18 data points; 1996/97)
20. ZEUS $F_2^c$ e$^+$p (27 data points; 1998/00)
21. H1 CC e$^+$p (28 data points; 1998-99)
22. H1 CC e$^+$p (25 data points; 1994-97)
23. H1 CC e$^+$p (28 data points; 1999-00)
24. ZEUS CC e$^+$p (26 data points; 1998-99)
25. ZEUS CC e$^+$p (29 data points; 1994-97)
26. ZEUS CC e$^+$p (30 data points; 1999-00)
27. NuTev neutrino dimuon cross section (38 data points)
28. NuTev anti-neutrino dimuon cross section (33 data points)
29. CCFR neutrino dimuon cross section (40 data points)
30. CCFR anti-neutrino cross section (38 data points)
31. E605 dimuon (199 data points)
32. E866 dimuon (13 data points)
33. Lepton asymmetry from CDF (11 data points)
34. CDF Run 1B jet cross section (33 data points)
35. D0 Run 1B jet cross section (90 data points)

- 2794 data points from DIS, DY, jet production
- All with (correlated) systematic errors that must be treated correctly in the fit
- Note that DIS is the 800 pound gorilla of the global fit with many data points and small statistical and systematic errors
  - and fixed target DIS data still have a significant impact on the global fitting, even with an abundance of HERA data
- To avoid non-perturbative effects, kinematic cuts on placed on the DIS data
  - $Q^2>5$ GeV$^2$
  - $W^2(=m^2+Q^2(1-x)/x)>12.25$ GeV$^2$
Influence of data in global fit

- Charged lepton DIS

\[ F_2(x, Q^2) = x \sum_i e_i^2 [q_i(x, Q^2) + \bar{q}_i(x, Q^2)] \]

- each flavor weighted by its squared charge
- quarks and anti-quarks enter together
- gluon doesn’t enter, in lowest order, but does enter into the structure functions at NLO
- also enters through mixing in evolution equations so gluon contributes to the change of the structure functions as \( Q^2 \) increases
- at low values of \( x \)

\[ Q^2 \frac{dF_2}{dQ^2} \approx \frac{\alpha_s}{2\pi} \sum_i e_i^2 \int_0^1 \frac{dy}{y} P_{qg}(y) G\left(\frac{x}{y}, Q^2\right) \]

- \( Q^2 \) dependence at small \( x \) is driven directly by gluon pdf

- At low \( x \), structure functions increase with \( Q^2 \); at high \( x \) decrease
Influence of data in global fit

**Neutrino DIS**

\[
F_2(x, Q^2) = x \sum [q_i(x, Q^2) + \bar{q}_i(x, Q^2)]
\]

\[
xF_3(x, Q^2) = x \sum [q_i(x, Q^2) - \bar{q}_i(x, Q^2)]
\]

- additional (parity-violating) structure function allows the separation of quarks and antiquarks but not a complete flavor separation
- caveat: neutrino observables usually obtained using nuclear targets so there is added question of nuclear corrections
Some observations from DIS

- DIS data provide strong constraints on the u and d distributions over the full range of x covered by the data.
- The combination $4\bar{u} + \bar{d}$ is well-constrained at small x.
- The gluon is constrained at low values of x by the slope of the $Q^2$ dependence of $F_2$.
- Momentum sum rule connects low x and high x behavior, but loosely.
Inclusive jets and global fits

- We don’t have many handles on the high $x$ gluon distribution in the global pdf fits
- Best handle is provided by the inclusive jet cross section from the Tevatron

At high $E_T$ (high $x$), $gq$ is subdominant, but there’s a great deal of freedom/uncertainty on the high $x$ gluon distribution

- about 42% of the proton’s momentum is carried by gluons, and most of that momentum is at low $x$

The inclusion of the CDF/D0 inclusive jet cross sections from Run 1 boosted the high $x$ gluon distribution and thus the predictions for the high $E_T$ jet cross sections

The high $x$ gluon has decreased due to influence of the Run 2 jet data

---

Figure 5.6. The subprocess contributions to inclusive jet production at the Tevatron for the CTEQ6M and CTEQ6M pdfs. The impact of the larger larger gluon at high $x$ for CTEQ6 is evident.
Global fitting: best fit

- Using our 2794 data points, we do our global fit by performing a $\chi^2$ minimization
  - where $D_i$ are the data points and $T_i$ are the theoretical predictions; we allow for a normalization shift $f_N$ for each experimental data set
    - but we provide a quadratic penalty for any normalization shift
  - where there are $k$ systematic errors $\beta$ for each data point in a particular data set
    - and where we allow the data points to be shifted by the systematic errors with the shifts given by the $s_j$ parameters
    - but we give a quadratic penalty for non-zero values of the shifts $s_j$
  - where $\sigma_i$ is the statistical error for data point $i$

- For each data set, we calculate

  $\chi^2 = \sum_i \left[ \frac{\left( f_N D_i - \sum_{j=1}^{k} \beta_{ij} s_j \right) - T_i}{\sigma_i^2} \right]^2 + \sum_{j=1}^{k} s_j^2$

- For a set of theory parameters it is possible to analytically solve for the shifts $s_j$, and therefore, continually update them as the fit proceeds

- To make matters more complicated, we may give additional weights to some experiments due to the utility of the data in those experiments (i.e. NA-51), so we adjust the $\chi^2$ to be

  $\chi^2 = \sum_k w_k \chi_k^2 + \sum_k w_{N,k} \left[ \frac{1 - f_N}{\sigma_N^{\text{norm}}} \right]^2$

- where $w_k$ is a weight given to the experimental data and $w_{N,k}$ is a weight given to the normalization
Minimization and errors

- Free parameters in the fit are parameters for quark and gluon distributions
  \[ f(x) = x^{(a_1 - 1)} (1 - x)^{a_2} e^{a_3 x} [1 + e^{a_4 x}]^{a_5} \]
- Too many parameters to allow all to remain free
  - some are fixed at reasonable values or determined by sum rules
- 20 free parameters for CTEQ6.1, 22 for CTEQ6.6, 24 for CT09, 26 for CT10

- Result is a global $\chi^2$/dof on the order of 1
  - for a NLO fit
  - worse for a LO fit, since the LO pdf’s can not make up for the deficiencies in the LO matrix elements
PDF Errors: old way

- Make plots of lots of pdf’s (no matter how old) and take spread as a measure of the error
- Can either underestimate or overestimate the error
- Review sources of uncertainty on pdf’s
  - data set choice
  - kinematic cuts
  - parametrization choices
  - treatment of heavy quarks
  - order of perturbation theory
  - errors on the data
- There are now more sophisticated techniques to deal with at least the errors due to the experimental data uncertainties
PDF Errors: new way

- 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

$$(\Delta X)^2 = \Delta \chi^2 \sum_{\mu, \nu} \frac{\partial X}{\partial a_\mu} (H^{-1})_{\mu \nu} \frac{\partial X}{\partial a_\nu}$$
PDF Errors: new way

- 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

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

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

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

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

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.
PDF Errors: new way

- 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

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

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

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.
Parametrization bias

- 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 bias) even more important.
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

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.
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

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

This works exactly for fixed order calculations and works well enough (see later) 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

On-line Plotting and Calculation.

Parton Distributions:

Using the form below you can calculate, in real time, values of $x f(x, Q^2)$ for any of the PDFs from the groups CTEQ, MRST, GRV/GJR, Alekhin, ZEUS and H1. You can also generate and compare plots of $x f(x)$ at any $Q^2$ for up to 4 different parton types or PDFs.

Parton Distributions with Error Analyses:

Select below if you wish the comparison of another PDF set with the above (note: this option only works for specific partons - not "all")

The CTEQ, MRST and ZEUS errors are calculated from the error analyses as described in their respective papers [link 1], [link 2], [link 3], and [link 4]. By summing over the pdfs given in the 40 (CTEQ), 30 (MRST) or 22 (ZEUS) eigenvector grids, in the following way:

The Alekhin errors are generated from quadratic summing of the derivatives of the pdfs over all the 15 parameters, as described in the fortran programme.
Let’s try it out

Up and down quarks dominate at high $x$, gluon at low $x$. As $Q^2$ increases, note the growth of the gluon distribution, and to a lesser extent the sea quark distributions.
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

\[ Q^2 = 2 \text{ GeV}^2 \]
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
  \[ \cos \varphi = \frac{\vec{\nabla} X \cdot \vec{\nabla} Y}{\Delta X \Delta Y} = \frac{1}{4 \Delta X \Delta Y} \sum_{i=1}^{N} \left( X_i^{(+)} - X_i^{(-)} \right) \left( Y_i^{(+)} - Y_i^{(-)} \right) \]
- The ellipse itself is given by
  \[ \left( \frac{\delta X}{\Delta X} \right)^2 + \left( \frac{\delta Y}{\Delta Y} \right)^2 - 2 \left( \frac{\delta X}{\Delta X} \right) \left( \frac{\delta Y}{\Delta Y} \right) \cos \varphi = \sin^2 \varphi \]

- If two cross sections/pdf's are very correlated, then $\cos \varphi \approx 1$
- …uncorrelated, then $\cos \varphi \approx 0$
- …anti-correlated, then $\cos \varphi \approx -1$
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
  \[ \cos \phi = \frac{\vec{V}_X \cdot \vec{V}_Y}{\Delta X \Delta Y} = \frac{1}{4\Delta X \Delta Y} \sum_{i=1}^N (X_i^{(+)} - X_i^{(-)}) (Y_i^{(+)} - Y_i^{(-)}) \]
- 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 $\cos \phi \approx -1$
Correlations between pdf’s

Correlations between $f(x_1, Q)$ and $f(x_2, Q)$ at $Q = 85$ GeV

Can you guess which PDF’s these are?

Homework assignment: which pdf’s and why?
Correlations between $f_1(x_1, Q)$ and $f_2(x_2, Q)$ at $Q = 85$ GeV

$d$ vs $u$

$s$ vs $\bar{u}$ at $Q=2$

$s$ vs. $g$

Sometimes there is a clear physics reason behind the correlation (e.g., sum rules or assumed Regge-like behavior); sometimes not
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]

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

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

for each of the five fixed values of $\alpha_s$. Then the overall best-fit prediction is $F^{\alpha_S}(S_0)$, $\alpha^{\theta}_S$ is the best-fit $\alpha_S$ value, and the overall “PDF+$\alpha_s$” uncertainties are given by

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

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

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)
  - $\alpha_s(m_Z) = 0.1184 +/- 0.0007$

- 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 +/-0.002 is a reasonable estimate of the uncertainty

G. Watt Mar 26 PDF4LHC meeting
\( \alpha_s(m_Z) \) 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
My recommendation to PDF4LHC/Higgs working group

- 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
  - (my suggestion) as standard, use 90%CL limits
  - 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
- In lecture #5
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
For CTEQ: $\alpha_s$ series

- Take CTEQ6.6 as base, and vary $\alpha_s(m_Z) \pm 0.002$ (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

This also means that one can naively scale between 68% and 90% CL.

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)
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

\[ \frac{\sigma(W^+)}{\sigma(W^-)} \text{ vs. } y_W \text{ at the LHC} \]

\[ \frac{\sigma(W^\pm)}{\sigma(Z^0)} \text{ vs. } y_{W/Z} \text{ at the LHC} \]
LO PDFs

- 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.
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 CTEQ6.5M 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