Uncertainties from Theory

M. Diehl

Deutsches Elektronen-Synchroton DESY

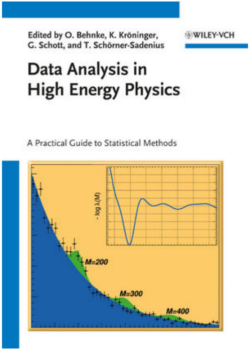
Statistics School 2014
Hamburg, 4 April 2014

HELMHOLTZ
GEMEINSCHAFT

DESY

1. Types of uncertainties
2. Perturbation theory and beyond
3. Uncertainties in parton density fits
4. Summary

more detail in chapter 9 of



Different sources of theoretical uncertainties

"observable = theoretical expression"

1. theoretical expression is only approximate
   - often obtained by expansion in small parameter
     e.g. in coupling constant → perturbation theory
     ⇝ estimate size of uncalculated/neglected terms
   - for some situations/aspects do not have systematic theory
     must use models
     ⇝ may estimate uncertainty by comparing different models

Different sources of theoretical uncertainties

"observable = theoretical expression"

2. input parameters from standard model, e.g.
   α_s, $m_{c,b}$, m_t, $m_{W,Z}$, m_H, CKM matrix elements
   note: running $\alpha_s(\mu)$ depends implicitly on quark masses
3. nonperturbative QCD parameters or functions
   - most prominently: parton distributions (PDFs)
     note: PDFs depend on $\alpha_s(\mu)$ via evolution
   - other examples: decay constants, wave functions
     (e.g. for $B \to D\ell\nu$, $B \to \pi K$)

quantities in points 2. and 3. may be obtained from

- comparison "measured observable = theor. expression"
- nonperturbative calculation (e.g. in lattice QCD)

Some parameters and their uncertainties

from: Review of Particles Physics 2012

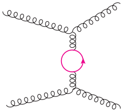
Phys. Rev. D86 (2012) 010001, http://hepdata.cedar.ac.uk/lbl

	relative uncertainty
$m_W = 80.385 \pm 0.015\,\mathrm{GeV}$	1.9×10^{-4}
$m_Z = 91.1876 \pm 0.0021\,\mathrm{GeV}$	2.3×10^{-5}
$m_\tau = 1.77682 \pm 0.00016\,\mathrm{GeV}$	9.0×10^{-5}
$m_b = 4.18 \pm 0.03\,\mathrm{GeV}$	0.72%
$m_t = 160^{+5}_{-4}\,\mathrm{GeV}$	$^{+3.1\%}_{-2.5\%}$
$1/\alpha_{em} = 137.035999074(44)$	3.2×10^{-9}
$\alpha_s(m_Z) = 0.1184 \pm 0.0007$	0.59%
$= 0.1183 \pm 0.0012$ (without lattice QCD)	1.0%

masses and couplings often appear raised to some power → larger uncertainties

m_b and m_t are $\overline{\mathrm{MS}}$ masses at scales $m_b(m_b)$, $m_t(m_t)$

Coupling constants and their running



- due to quantum effects (loop corrections)
  coupling constants depend on renormalization scale μ
  set by "typical momentum scale" of physical process
- electroweak interactions: $\alpha_{em}(0) \approx 0.00730$
  generally quick convergence of pert. series
- strong interactions:
  $\alpha_s(m_Z) \approx 0.118, \quad \alpha_s(m_\tau) \approx 0.33$
  higher order corrections often
  very important (easily factors of 2)
  coupling grows with decreasing μ
  $\rightsquigarrow$ expansion in α_s not useful at low μ

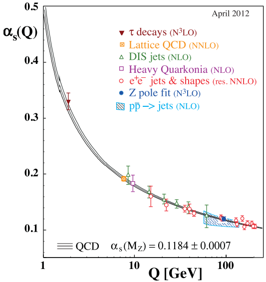


figure: Rev. Part. Phys. 2012

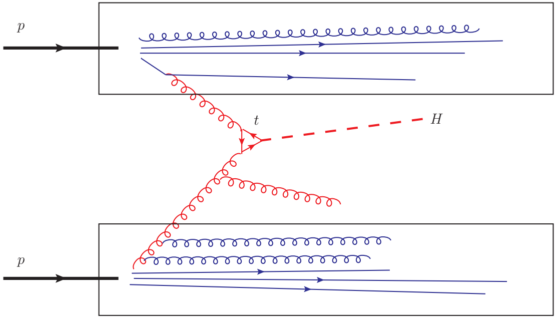
- not everything can be expanded in coupling (e.g. $e^{-\text{const}/\alpha}$)
  non-perturbative effects most ubiquitous in strong interactions

will concentrate on uncertainties due to strong interactions

Factorization: a cornerstone of calculations in QCD

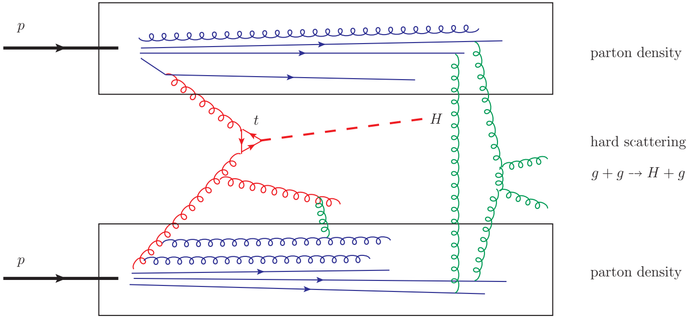
- confinement: quarks and gluons do not exist as free particles; only hadrons are observed
  $\rightsquigarrow$ even at very high energies pp (and ep) collisions involve dynamics at scales $\sim 1\,\mathrm{GeV}$ and below
- idea: separate physics at high and low momentum scales
  - at high scales use expansion in α_s
  - at low scales:
    determine non-perturbative quantities (e.g. parton densities) from experiment, theory or models
    once they are determined, we have predictive power

Example: $p + p \to H + X$ ($X =$ anything)



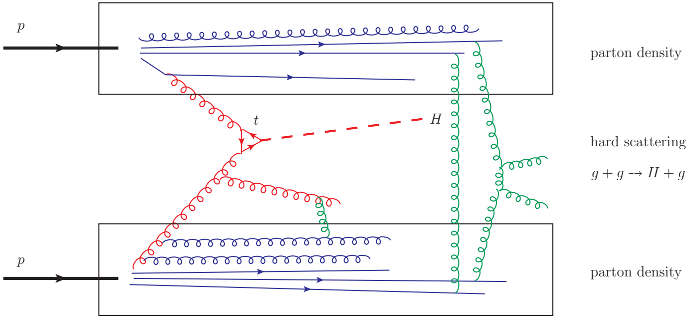
- factorization formula: parton densities and hard-scattering subprocesses $gg \to H$, $gg \to H + g$, ...

Example: $p + p \to H + X$ ($X =$ anything)



- actual physics more complicated: soft gluon exchange
  - outside domain of perturbation theory $\rightsquigarrow$ must model
  - does not affect sufficiently inclusive observables
  - but does matter for details of final state $\rightsquigarrow$ "underlying event"

Example: $p + p \rightarrow H + X$ (X = anything)



- if additional interactions hard ⇝ "multiparton interactions"
  suppressed in sufficiently inclusive observables
  no systematic theory yet ⇝ must model

Factorization formulae

schematic structure for pp collisions

$$\frac{d\sigma}{d(\text{variables})} = f_1(\mu_F) \underset{x_1}{\otimes} \frac{1}{Q^n}\, C\Big\{\frac{\mu_F}{Q}, \frac{\mu_R}{Q}, \alpha_s(\mu_R), \ldots\Big\} \underset{x_2}{\otimes} f_2(\mu_F)$$
$$+\, \mathcal{O}\Big(\frac{1}{Q^{n+1}} \text{ or } \frac{1}{Q^{n+2}}\Big)$$

- $f_1, f_2 =$ parton densities, $\quad C =$ hard-scattering coefficient
- $Q =$ hard momentum scale (e.g. Higgs mass, jet E_T)
  $x_1, x_2 =$ dimensionless variables constructed from kinematics
- convolution $f \underset{x}{\otimes} g = \int_x^1 \frac{dz}{z} f\Big(\frac{x}{z}\Big)\, g(z)$
- in $C(\ldots)$ possible dependence on m_t, m_H etc.
- higher-order corrections (1st line) $\rightarrow$ next slides
  power corrections (2nd line) $\rightarrow$ not discussed here

Factorization formulae

schematic structure for pp collisions

$$\frac{d\sigma}{d(\text{variables})} = f_1(\mu_F) \underset{x_1}{\otimes} \frac{1}{Q^n}\, C\Big\{ \frac{\mu_F}{Q}, \frac{\mu_R}{Q}, \alpha_s(\mu_R), \ldots \Big\} \underset{x_2}{\otimes} f_2(\mu_F)$$
$$+ \mathcal{O}\Big(\frac{1}{Q^{n+1}} \text{ or } \frac{1}{Q^{n+2}} \Big)$$

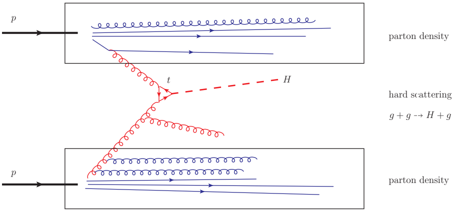
- have α_s expansions for C and for $df/d\mu_F$
- $\mu_R =$ renormalization scale
  separates physics at scale Q from physics at much higher scales (ultraviolet region)

Factorization formulae

schematic structure for pp collisions

$$\frac{d\sigma}{d(\text{variables})} = f_1(\mu_F) \underset{x_1}{\otimes} \frac{1}{Q^n}\, C\Big\{ \frac{\mu_F}{Q}, \frac{\mu_R}{Q}, \alpha_s(\mu_R), \ldots \Big\} \underset{x_2}{\otimes} f_2(\mu_F) + \mathcal{O}\Big(\frac{1}{Q^{n+1}} \text{ or } \frac{1}{Q^{n+2}} \Big)$$

- have α_s expansions for C and for $df/d\mu_F$
- $\mu_F =$ factorization scale
  separates physics at scale Q from physics at lower scales



Renormalization scale dependence

on next slides write μ instead of μ_R for brevity

- renormalization group equation

$$\frac{d}{d\log\mu^2}\,\alpha_s(\mu) = \beta\big(\alpha_s(\mu)\big)$$

with $\beta(\alpha_s) = -\alpha_s^2 \big(b_0^{n_f} + b_1^{n_f}\alpha_s + b_2^{n_f}\alpha_s^2 + b_3^{n_f}\alpha_s^3 + \ldots\big)$

- in practice: truncate series of $\beta(\alpha_s)$ and solve RGE numerically or analytically (possibly approximate)
- higher coefficients in α_s expansion of hard-scattering coefficient are μ dependent

$$C = \alpha_s^m(\mu)\,C_0 + \alpha_s^{m+1}(\mu)\,C_1\big(\tfrac{Q}{\mu}\big) + \alpha_s^{m+2}(\mu)\,C_2\big(\tfrac{Q}{\mu}\big) + \ldots$$

but C is independent of μ to any given accuracy in α_s:

$$\frac{d}{d\log\mu^2}\,C(\mu) = 0$$

see how this works:

- set $\mu = Q$ in expansion:

$$C = \alpha_s^m(\mu)\, C_0 + \alpha_s^{m+1}(\mu)\, C_1\big(\tfrac{Q}{\mu}\big) + \alpha_s^{m+2}(\mu)\, C_2\big(\tfrac{Q}{\mu}\big) + \dots$$
$$= \alpha_s^m(Q)\, C_0 + \alpha_s^{m+1}(Q)\, C_1(1) + \alpha_s^{m+2}(Q)\, C_2(1) + \dots$$

- expand $\alpha_s(Q) = \alpha_s(\mu) + a_1\big(\tfrac{Q}{\mu}\big)\, \alpha_s^2(\mu) + a_2\big(\tfrac{Q}{\mu}\big)\, \alpha_s^3(\mu) + \mathcal{O}(\alpha_s^4)$

see how this works:

- set $\mu = Q$ in expansion:

$$C = \alpha_s^m(\mu)\, C_0 + \alpha_s^{m+1}(\mu)\, C_1\big(\tfrac{Q}{\mu}\big) + \alpha_s^{m+2}(\mu)\, C_2\big(\tfrac{Q}{\mu}\big) + \ldots$$
$$= \alpha_s^m(Q)\, C_0 + \alpha_s^{m+1}(Q)\, C_1(1) + \alpha_s^{m+2}(Q)\, C_2(1) + \ldots$$

- expand $\alpha_s(Q) = \alpha_s(\mu) + a_1\big(\tfrac{Q}{\mu}\big)\, \alpha_s^2(\mu) + a_2\big(\tfrac{Q}{\mu}\big)\, \alpha_s^3(\mu) + \mathcal{O}(\alpha_s^4)$

$$\frac{d}{d\log Q^2}\,\text{(l.h.s.)} = \beta\big(\alpha_s(Q)\big) = -b_0\, \alpha_s^2(Q) - b_1\, \alpha_s^3(Q) + \mathcal{O}(\alpha_s^4)$$
$$= -b_0\, \alpha_s^2(\mu) - 2a_1 b_0\, \alpha_s^3(\mu) - b_1\, \alpha_s^3(\mu) + \mathcal{O}(\alpha_s^4)$$
$$\frac{d}{d\log Q^2}\,\text{(r.h.s.)} = \frac{da_1}{d\log Q^2}\, \alpha_s^2(\mu) + \frac{da_2}{d\log Q^2}\, \alpha_s^3(\mu) + \mathcal{O}(\alpha_s^4)$$

see how this works:

- set $\mu = Q$ in expansion:

$$C = \alpha_s^m(\mu)\, C_0 + \alpha_s^{m+1}(\mu)\, C_1\big(\tfrac{Q}{\mu}\big) + \alpha_s^{m+2}(\mu)\, C_2\big(\tfrac{Q}{\mu}\big) + \ldots$$
$$= \alpha_s^m(Q)\, C_0 + \alpha_s^{m+1}(Q)\, C_1(1) + \alpha_s^{m+2}(Q)\, C_2(1) + \ldots$$

- expand $\alpha_s(Q) = \alpha_s(\mu) + a_1\big(\tfrac{Q}{\mu}\big)\, \alpha_s^2(\mu) + a_2\big(\tfrac{Q}{\mu}\big)\, \alpha_s^3(\mu) + \mathcal{O}(\alpha_s^4)$

$$\frac{d}{d\log Q^2}\,(\text{l.h.s.}) = \beta\big(\alpha_s(Q)\big) = -b_0\, \alpha_s^2(Q) - b_1\, \alpha_s^3(Q) + \mathcal{O}(\alpha_s^4)$$
$$= -b_0\, \alpha_s^2(\mu) - 2a_1 b_0\, \alpha_s^3(\mu) - b_1\, \alpha_s^3(\mu) + \mathcal{O}(\alpha_s^4)$$
$$\frac{d}{d\log Q^2}\,(\text{r.h.s.}) = \frac{da_1}{d\log Q^2}\, \alpha_s^2(\mu) + \frac{da_2}{d\log Q^2}\, \alpha_s^3(\mu) + \mathcal{O}(\alpha_s^4)$$

- compare coefficients of $\alpha_s^n(\mu)$:

$$\frac{da_1}{d\log Q^2} = -b_0 \qquad \Rightarrow \qquad a_1\big(\tfrac{Q}{\mu}\big) = -b_0\, \log \tfrac{Q^2}{\mu^2}$$
$$\frac{da_2}{d\log Q^2} = -2a_1 b_0 - b_1 \qquad \Rightarrow \qquad a_2\big(\tfrac{Q}{\mu}\big) = +b_0^2\, \log^2 \tfrac{Q^2}{\mu^2} - b_1\, \log \tfrac{Q^2}{\mu^2}$$

- inserting

$$\alpha_s(Q) = \alpha_s(\mu)\Big[1-\alpha_s(\mu)\, b_0 \log\tfrac{Q^2}{\mu^2} + \alpha_s^2(\mu)\Big(b_0^2 \log^2\tfrac{Q^2}{\mu^2} - b_1 \log\tfrac{Q^2}{\mu^2}\Big) + \ldots\Big]$$

into $C = \alpha_s^m(Q)\Big[C_0 + \alpha_s(Q)\, C_1(1) + \alpha_s^2(Q)\, C_2(1) + \ldots\Big]$ get

$$\begin{aligned} C &= \alpha_s^m(\mu) \\ &\quad \times \Big[1 - \alpha_s(\mu)\, m b_0 \log\tfrac{Q^2}{\mu^2} + \alpha_s^2(\mu)\Big(\tfrac{m(m+1)}{2}\, b_0^2 \log^2\tfrac{Q^2}{\mu^2} - m b_1 \log\tfrac{Q^2}{\mu^2}\Big)\Big] \\ &\quad \times \Big[C_0 + \alpha_s(\mu)\, C_1(1) + \alpha_s^2(\mu)\Big(C_2(1) - C_1(1)\, b_0 \log\tfrac{Q^2}{\mu^2}\Big)\Big] + \mathcal{O}(\alpha_s^{m+3}) \end{aligned}$$

- inserting

$$\alpha_s(Q) = \alpha_s(\mu)\Big[1-\alpha_s(\mu)\, b_0 \log\tfrac{Q^2}{\mu^2} + \alpha_s^2(\mu)\Big(b_0^2 \log^2\tfrac{Q^2}{\mu^2} - b_1 \log\tfrac{Q^2}{\mu^2}\Big) + \ldots\Big]$$

into $\quad C = \alpha_s^m(Q)\, \Big[C_0 + \alpha_s(Q)\, C_1(1) + \alpha_s^2(Q)\, C_2(1) + \ldots\Big]$ get

$$\begin{aligned} C = {} & \alpha_s^m(\mu) \\ & \times \Big[1 - \alpha_s(\mu)\, m b_0 \log\tfrac{Q^2}{\mu^2} + \alpha_s^2(\mu)\Big(\tfrac{m(m+1)}{2}\, b_0^2 \log^2\tfrac{Q^2}{\mu^2} - m b_1 \log\tfrac{Q^2}{\mu^2}\Big)\Big] \\ & \times \Big[C_0 + \alpha_s(\mu)\, C_1(1) + \alpha_s^2(\mu)\,\Big(C_2(1) - C_1(1)\, b_0 \log\tfrac{Q^2}{\mu^2}\Big)\Big] + \mathcal{O}(\alpha_s^{m+3}) \end{aligned}$$

- in $\quad C = \alpha_s^m(\mu)\, C_0 + \alpha_s^{m+1}(\mu)\, C_1\big(\tfrac{Q}{\mu}\big) + \alpha_s^{m+2}(\mu)\, C_2\big(\tfrac{Q}{\mu}\big) + \ldots$
  have coefficients

$$C_1\big(\tfrac{Q}{\mu}\big) = C_1(1) - m b_0\, C_0 \log\tfrac{Q^2}{\mu^2}$$

$$C_2\big(\tfrac{Q}{\mu}\big) = C_2(1) - \Big[(m+1) b_0\, C_1(1) + m b_1\, C_0\Big] \log\tfrac{Q^2}{\mu^2} + \tfrac{m(m+1)}{2}\, b_0^2\, C_0 \log^2\tfrac{Q^2}{\mu^2}$$

- inserting

$$\alpha_s(Q) = \alpha_s(\mu)\Big[1-\alpha_s(\mu)\, b_0 \log\tfrac{Q^2}{\mu^2} + \alpha_s^2(\mu)\Big(b_0^2 \log^2\tfrac{Q^2}{\mu^2} - b_1 \log\tfrac{Q^2}{\mu^2}\Big) + \ldots\Big]$$

into $\quad C = \alpha_s^m(Q)\,\Big[C_0 + \alpha_s(Q)\, C_1(1) + \alpha_s^2(Q)\, C_2(1) + \ldots\Big]$ get

$$\begin{aligned} C &= \alpha_s^m(\mu) \\ &\quad\times \Big[1 - \alpha_s(\mu)\, m b_0 \log\tfrac{Q^2}{\mu^2} + \alpha_s^2(\mu)\Big(\tfrac{m(m+1)}{2}\, b_0^2 \log^2\tfrac{Q^2}{\mu^2} - m b_1 \log\tfrac{Q^2}{\mu^2}\Big)\Big] \\ &\quad\times \Big[C_0 + \alpha_s(\mu)\, C_1(1) + \alpha_s^2(\mu)\,\Big(C_2(1) - C_1(1)\, b_0 \log\tfrac{Q^2}{\mu^2}\Big)\Big] + \mathcal{O}(\alpha_s^{m+3}) \end{aligned}$$

- in $\quad C = \alpha_s^m(\mu)\, C_0 + \alpha_s^{m+1}(\mu)\, C_1\big(\tfrac{Q}{\mu}\big) + \alpha_s^{m+2}(\mu)\, C_2\big(\tfrac{Q}{\mu}\big) + \ldots$
  have coefficients

$$\begin{aligned} C_1\big(\tfrac{Q}{\mu}\big) &= C_1(1) - m b_0\, C_0 \log\tfrac{Q^2}{\mu^2} \\ C_2\big(\tfrac{Q}{\mu}\big) &= C_2(1) - \Big[(m+1) b_0\, C_1(1) + m b_1\, C_0\Big] \log\tfrac{Q^2}{\mu^2} + \tfrac{m(m+1)}{2}\, b_0^2\, C_0 \log^2\tfrac{Q^2}{\mu^2} \end{aligned}$$

- check (exercise): $\frac{d}{d\log\mu^2} C\big(\tfrac{Q}{\mu}, \alpha_s(\mu)\big) = \Big[\frac{\partial}{\partial\log\mu^2} + \beta\,\frac{\partial}{\partial\alpha_s}\Big] C = 0$

- have

$$C = \alpha_s^m(\mu)\, C_0 + \alpha_s^{m+1}(\mu)\, C_1\big(\tfrac{Q}{\mu}\big) + \alpha_s^{m+2}(\mu)\, C_2\big(\tfrac{Q}{\mu}\big) + \ldots$$

with

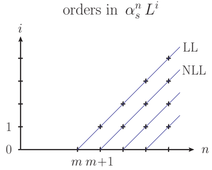
$$C_1\big(\tfrac{Q}{\mu}\big) = C_1(1) - m b_0 C_0 \log \tfrac{Q^2}{\mu^2}$$

$$C_2\big(\tfrac{Q}{\mu}\big) = C_2(1) - \Big[(m+1) b_0 C_1(1) + m b_1 C_0\Big] \log \tfrac{Q^2}{\mu^2} + \tfrac{m(m+1)}{2} b_0^2 C_0 \log^2 \tfrac{Q^2}{\mu^2}$$

- calculating C_0 (LO) get also terms $\alpha_s^{m+1} \log \frac{Q^2}{\mu^2}, \alpha_s^{m+2} \log^2 \frac{Q^2}{\mu^2}, \ldots$
  calculating $C_1(1)$ (NLO) get also terms $\alpha_s^{m+2} \log \frac{Q^2}{\mu^2}, \alpha_s^{m+3} \log^2 \frac{Q^2}{\mu^2}, \ldots$
  $\rightsquigarrow$ recover logarithmic terms at higher orders, but not coefficients $C_n(1)$
- varying μ in N^{n}LO result get variation at N^{n+1}LO corresponding to

$$\alpha_s^{n+1} \sum_{i=1}^{n+1-m} (\text{known coeff.}) \times \log^i \frac{\mu^2}{Q^2} + \mathcal{O}(\alpha_s^{n+2})$$

  but no information on $\alpha_s^{n+1} C_{n+1}(1)$



Renormalization scale dependence

- varying μ in N^{n}LO result get variation at N^{n+1}LO corresponding to
  $$\alpha_s^{n+1} \sum_{i=1}^{n+1-m} (\text{known coeff.}) \times \log^i \frac{\mu^2}{Q^2} + \mathcal{O}(\alpha_s^{n+2})$$
  but no information on $\alpha_s^{n+1} C_{n+1}(1)$

consequences:

- when calculate higher orders
  expect that scale dependence decreases
- scale variation in N^{n}LO result estimates size of certain higher-order terms, but not of all
  - uncalculated higher orders often estimated by varying μ between $1/2$ and 2 times some central value
    is a conventional choice
  - but what to take for central value?

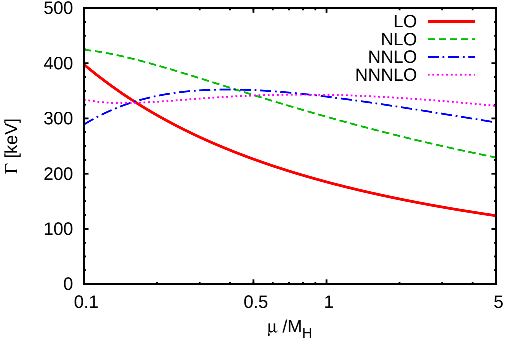
Renormalization scale choice

- prescriptions for scale choice aiming to minimizing size of higher-order terms

  take NLO calc. of $C(\mu) = \alpha_s^m C_0 + \alpha_s^{m+1} C_1(\mu) + \mathcal{O}(\alpha_s^{m+2})$
  - $\mu =$ typical virtuality in hard-scattering graphs
    useful guidance, but obviously not a well-defined quantity
  - principle of minimal sensitivity (PMS): $\frac{d}{d\mu^2} \sum_{i=0}^{1} \alpha_s^{m+i} C_i(\mu) = 0$
  - fastest apparent convergence (FAC): $C_1(\mu) = 0$
  - Brodsky-Mackenzie-Lepage (BLM): more complicated
- how much these reduce higher orders depends on process
  cannot “predict” higher orders without calculating them

Renormalization scale dependence

- example: inclusive hadronic decay of Higgs boson via top quark loop (i.e. without direct coupling to $b\bar{b}$)
- in perturbation theory: $H \to 2g$, $H \to 3g$, ... known to N^{3}LO Baikov, Chetyrkin, hep-ph/0604194



plot for $m_H = 125\,\mathrm{GeV}$

Factorization scale dependence

- scale dependence of PDF given by DGLAP equation:

$$\frac{d}{d\log\mu_F^2}\,\mathsf{PDF}(x,\mu_F) = \mathsf{PDF}(\mu_F) \underset{x}{\otimes} P\big(\alpha_s(\mu_F)\big)$$

  evolution kernels have perturbative expansion in α_s :

$$P\big(z,\alpha_s(\mu_F)\big) = \alpha_s(\mu_F)\,P_0(z) + \alpha_s^2(\mu_F)\,P_1(z) + \mathcal{O}(\alpha_s^3)$$

  - choose approx. of evolution kernel (LO, NLO, NNLO)
  - solve DGLAP equations numerically
    $\Rightarrow$ obtain PDF(μ_1) from PDF(μ_0)

- hard-scattering coefficient contains powers of $\log(\mu_F/Q)$
  μ_F independence of PDF$(\mu_F) \otimes C(\mu_F)$ implies

$$\frac{d}{d\log\mu_F^2}\,C\big(x,\mu_F,\mu_R,\alpha_s(\mu_R)\big) = -P\big(\alpha_s(\mu_F)\big) \underset{x}{\otimes} C\big(\mu_F,\mu_R,\alpha_s(\mu_R)\big)$$

Factorization scale dependence

$$\frac{d}{d\log\mu_F^2}\, C\big(x,\mu_F,\mu_R,\alpha_s(\mu_R),\ldots\big) = -P\big(\alpha_s(\mu_F)\big) \underset{x}{\otimes} C\big(\mu_F,\mu_R,\alpha_s(\mu_R)\big)$$

using renormalization group equation can rewrite

$$\alpha_s(\mu_F) = \alpha_s(\mu_R) + a_1\big(\tfrac{\mu_F}{\mu_R}\big)\, \alpha_s^2(\mu_R) + \mathcal{O}(\alpha_s^3)$$

with expansions

$$\begin{aligned} P\big(\alpha_s(\mu_F)\big) &= P_0\,\alpha_s(\mu_F) + P_1\,\alpha_s^2(\mu_F) + \mathcal{O}(\alpha_s^3) \\ &= P_0\,\alpha_s(\mu_R) + \mathcal{O}(\alpha_s^2) \\ C\big(\mu_F,\mu_R,\alpha_s(\mu_R)\big) &= C_0 + C_1(\mu_F,\mu_R)\,\alpha_s(\mu_R) + \mathcal{O}(\alpha_s^2) \end{aligned}$$

can match coefficients order by order

$$\frac{d}{d\log\mu_F^2}\, C_1(\mu_F,\mu_R) = -P_0\otimes C_0 \;\;\Rightarrow\;\; C_1(\mu_F,\mu_R) = C_1(Q,\mu_R) - P_0\otimes C_0\;\log\frac{\mu_F^2}{Q^2}$$

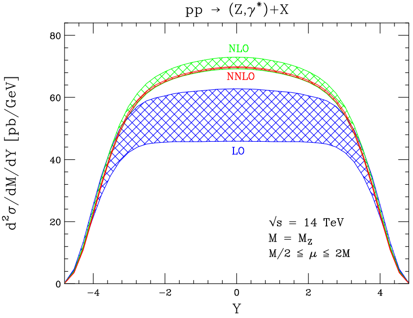
Factorization scale dependence

- try to chose μ_F such as to avoid large higher-order coefficients
- with C calculated to N^{n}LO have μ_F dependence of order N^{n+1}LO in convolution PDF $\otimes\, C$

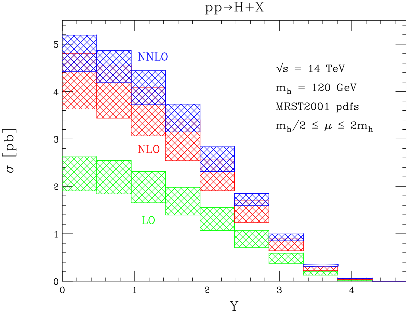
  if evolve PDFs with DGLAP kernels up to $\alpha_s^n \, P_{n-1}$ or higher
- as for μ_R may estimate certain higher-order terms by varying μ_F between e.g. $1/2$ and 2 times some central value
- as for μ_R no general solution for finding μ_F that minimizes higher orders
- often set $\mu_F = \mu_R$ (and vary them) together
  but can also set and vary them separately

Scale dependence

examples: rapidity distributions in Z/γ^* and in Higgs production



Anastasiou, Dixon, Melnikov, Petriello, hep-ph/0312266



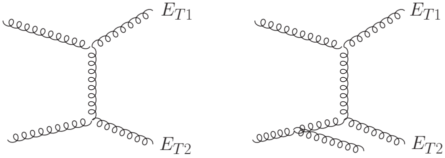
Anastasiou, Melnikov, Petriello, hep-ph/0501130

$\mu_F = \mu_R = \mu$ varied within factor 1/2 to 2

LO, NLO, and higher

- instead of varying scale(s) may estimate higher orders by comparing N^{n}LO result with N^{n-1}LO
- caveat: comparison NLO vs. LO may not be representative for situation at higher orders

  often have especially large step from LO to NLO
  - certain types of contribution may first appear at NLO
    e.g. terms with gluon density $g(x)$ in DIS, $pp \to W + X$, etc.
  - final state at LO may be too restrictive
    e.g. in $\frac{d\sigma}{dE_{T1}\, dE_{T2}}$ for dijet production



Multi-scale problems

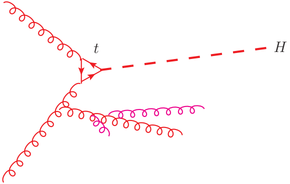
- scale choice even less obvious when have several hard scales
  e.g. Q and p_T, Q and m_c, p_T and m_W, ...
  may try to identify typical virtualities in graphs
- for small/large ratios of hard scales
  (or small/large values of scaling variables, e.g. $x \to 0$ or $x \to 1$)
  may have large logarithms in C for any choice of μ_R, μ_F

Multi-scale problems

- for certain cases can resum large logarithms to all orders
  e.g. $\alpha_s^n \log^{n+i}$ for all n with given $i = 0, 1, \ldots$
  - transverse-momentum logs: $\log \frac{p_T}{Q}$ for $p_T \ll Q$
    $\rightsquigarrow$ Sudakov factors
  - threshold logs: $\log \frac{M^2}{\hat{s}}$ for $\hat{s} \to M^2$
    for production of mass M with partonic collision energy $\sqrt{\hat{s}}$
    $\sigma(pp) \sim \int dz_1 dz_2 \, \text{PDF}(z_1) \, \text{PDF}(z_2) \, C(\hat{s} = z_1 z_2 s)$
  - high-energy logs: $\log \frac{Q^2}{s}$ for $s \gg Q^2$ $\rightsquigarrow$ BFKL equation
- resummation procedure may have its own uncertainties
  e.g. from integrals of type $\int_0^Q d\mu \, \text{fct.}\big(\alpha_s(\mu)\big)$ $\rightsquigarrow$ Landau pole

Jet production

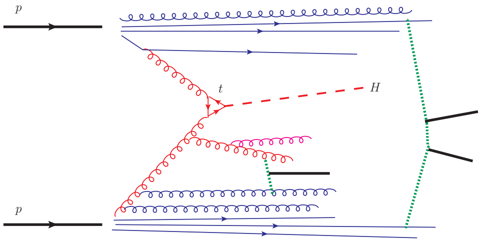
- fundamental problem: factorization formulae are for prod'n of high-p_T partons, not high-p_T hadrons
- parton $\rightarrow$ hadron transition non-perturbative $\rightsquigarrow$ need model
- to minimize theory uncertainties:
  - define hadronic jets using an algorithm that is not sensitive to collinear and soft radiation (beyond perturbative control)



  - apply algorithm to partons in computation and to hadrons in measurement
  - hadronization corrections should then be moderate and typically decrease with jet E_T

Monte Carlo generators e.g. HERWIG, PYTHIA, SHERPA

- build on structure of factorization formulae
- but compute fully exclusive final states (no unspecified "$+X$")



- ingredients:
  - parton densities and hard-scattering matrix elements
  - parton showers: small-angle radiation from partons in initial and final state in perturbative region
  - models for hadronization, underlying event, multiparton int's

Parton density fits

Principle of PDF determinations:

- compare data with factorization formulae for selected processes and kinematics
- specify PDF at reference scale μ_0

  use DGLAP eqs. to evolve to scales μ used in fact. formulae
- conventional determinations parameterize PDFs at μ_0 and determine parameters by χ^2 fit to data

  NNPDF collab. uses neural networks, avoids choice of function claims "unbiased" representation of PDFs

Recent PDF sets

PDF set	order	fitted PDF parameters	μ_0^2 [GeV^2]	Q^2_{min} [GeV^2]	$\alpha_s(m_Z)$	
JR09	NNLO	20	0.55	2	$0.1124(20)$	fitted
ABKM09	NNLO	21	9	2.5	$0.1135(14)$	fitted
MSTW08	LO	28	1	2	0.139	
	NLO				0.120	
	NNLO				0.117	
HERAPDF1.0	NLO	10	1.9	3.5	0.1176	
CT10	NLO	26	1.69	4	0.118	
NNPDF2.1	LO	259	2	3	$0.119, 0.130$	
	NLO				0.119	
	NNLO				0.119	

Q_{min} = minimum Q for fitted data on deep inelastic scattering
updates: ABKM09 → ABM12, HERAPDF1.0 → 1.5, NNPDF2.1 → 2.3

Recent PDF sets

PDF set	m_c [GeV]	m_b [GeV]	tolerance T 68% CL	tolerance T 90% CL
JR09	1.3	4.2	4.54	
ABKM09	1.5	4.5	1	
MSTW08	1.4	4.75	≈ 1 to 6.5	≈ 2.5 to 11
HERAPDF1.0	1.4	4.75	1	
CT10	1.3	4.75		10
NNPDF2.1	1.414	4.75	–	–

m_c and m_b are pole masses

T will be explained later

Uncertainties on extracted PDFs

"systematic theory uncertainties"

- selection of data sets and kinematics
- perturbative order of evolution and hard-scattering coefficients
- values of α_s and m_c, m_b and possibly other constants
  if taken as external parameters i.e. not fitted
  some PDF sets are available for different values of α_s
- fine details of perturbative calculations
  e.g. treatment of heavy quarks, resummation
- power corrections (typically try to avoid by minimal Q in data)
- corrections for data with nuclear targets

errors on fitted parameters

- reflect errors (stat. and syst.) of fitted data
  discuss on the following slides

Parametric errors in PDF fits

see e.g. hep-ph/0201195 (CTEQ6), arXiv:0802.0007 (CTEQ6.6)
arXiv:0901.0002 (MSTW 2008)

- errors obtained in χ^2 fit

simplest version: $$\chi^2 = \sum_i \frac{\big[D_i - T_i(\boldsymbol{p})\big]^2}{\sigma^2_{i,\,\mathsf{stat}} + \sigma^2_{i,\,\mathsf{syst}}}$$

$D_i =$ data point number i

$T_i =$ corresponding theory prediction

$\boldsymbol{p} = \{p_1, \ldots, p_k\} =$ set of fitting parameters

more sophisticated treatment for correlated systematic errors, i.e. overall normalization

$$\chi^2 = \sum_i \frac{\left[D_i - T_i(\boldsymbol{p})\right]^2}{\sigma^2_{i,\,\mathsf{stat}} + \sigma^2_{i,\,\mathsf{syst}}}$$

if assume that errors of D_i follow a Gaussian distribution, then

- parameters $\boldsymbol{p}_{\mathsf{min}}$ that minimize χ^2
  follow a k-dim. Gaussian dist. around true values $\boldsymbol{p}_0$
- $\Delta\chi^2(\boldsymbol{p}) = \chi^2(\boldsymbol{p}) - \chi^2_{\mathsf{min}} = \sum_{ij} (p - p_{\mathsf{min}})_i \, H_{ij} \, (p - p_{\mathsf{min}})_j$
  $H =$ Hesse matrix = inverse of covariance matrix V
- observable $\mathcal{O}(\boldsymbol{p}_{\mathsf{min}})$ follows Gaussian dist. with error

$$\Delta\mathcal{O} = T \sqrt{\sum_{ij} \frac{\partial\mathcal{O}}{\partial p_i} \, H^{-1}_{ij} \, \frac{\partial\mathcal{O}}{\partial p_j}}$$

with $T = 1$ for 68% C.L., $T = 2.71$ for 95% C.L. etc.
readily generalizes to several obs. and their correlated errors
$\rightsquigarrow$ complicated in practice, would need derivatives $\partial\mathcal{O}/\partial p_i$

$$\Delta\mathcal{O} = T\sqrt{\sum\nolimits_{ij} \frac{\partial \mathcal{O}}{\partial p_i}\, H^{-1}_{ij}\, \frac{\partial \mathcal{O}}{\partial p_j}}$$

- diagonalize Hesse matrix H and rescale eigenvectors
  $\Rightarrow$ linear combinations z_i of $(p - p_{\text{min}})_j$ satisfying

$$\Delta\chi^2 = \sum\nolimits_{ij} (p - p_{\text{min}})_i\, H_{ij}\, (p - p_{\text{min}})_j = \sum\nolimits_i z_i^2$$

$$\Delta\mathcal{O} = T\sqrt{\sum\nolimits_i \frac{\partial \mathcal{O}}{\partial z_i}\, \frac{\partial \mathcal{O}}{\partial z_i}} = \sqrt{\sum\nolimits_i \left[\frac{\mathcal{O}(S_i^+) - \mathcal{O}(S_i^-)}{2}\right]^2}$$

  with eigenvector PDF sets $S_i^\pm$
  corresponding to parameters $z_i = \pm T$ and $z_j = 0$ for $j \neq i$
  in last step have linearized $\mathcal{O}$ around $\boldsymbol{z} = 0$

- for large errors $\Delta\chi^2$ not quadratic in $(p - p_{\text{min}})_i$ or z_i
  $\rightsquigarrow$ linear error propagation not reliable
  $\rightsquigarrow$ Lagrange multiplier method (not discussed here)
  see e.g. CTEQ, hep-ph/0101051

The tolerance criterion

- if data points D_i follow Gaussian distribution then experiment with N_j data points contributes $\chi^2_{j,\text{min}} \sim N_j$ to global χ^2_{min}
- not always seen in practice: for some cases
  - $\chi^2_{j,\text{min}}$ significantly below or above N_j
  - $\chi^2_{j,\text{min}}$ much larger than χ^2 minimized separately for experiment

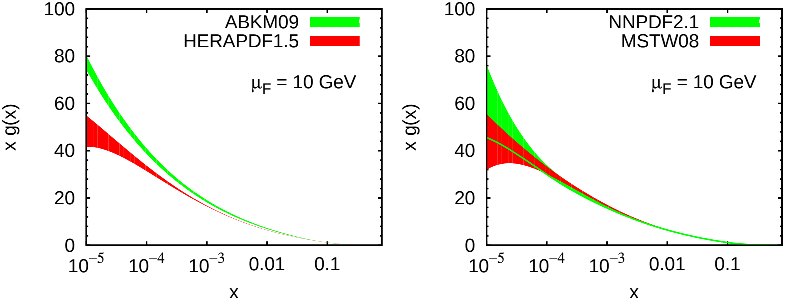
  may be due to inconsistent data sets, shortcomings of theory description or of PDF parameterization
  in such a case standard χ^2 errors misrepresent uncertainty
- modified criterion for T adopted by groups CT, MSTW, JR
  - obtained by procedure/algorithm looking at χ^2 from individual experiments
  - may be seen as ad hoc deviation from "standard statistics" but "standard criterion" for T requires that all data points have Gaussian dist. with quoted uncertainties

The NNPDF approach

- to avoid bias due to functional form of PDFs, use very flexible functions (called "neural networks") with large number of free parameters (about $10\times$ more than in standard approach)
- standard χ^2 fit then not possible, instead
  - generate Monte Carlo ensemble of replicas of original data (typically $N_{\text{rep}} = 100$ or 1000), according to central values and errors of measurement
  - in each replica divide data into "training" and "validation" set
  - for each replica minimize χ^2 on training set until χ^2 of validation set starts to increase (thus avoiding to fit "noise in the data")
- central values and uncertainties on observable (or on PDFs themselves) given by ensemble average etc.

$$\overline{\mathcal{O}} = \frac{1}{N_{\text{rep}}} \sum_{r}^{N_{\text{rep}}} \mathcal{O}_r \qquad (\Delta\mathcal{O})^2 = \frac{1}{N_{\text{rep}} - 1} \sum_{r}^{N_{\text{rep}}} (\mathcal{O}_r - \overline{\mathcal{O}})^2$$

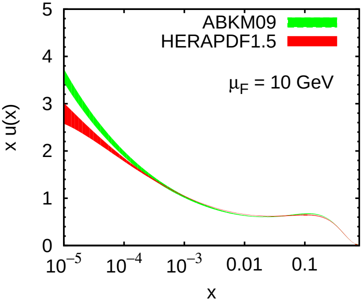
Illustration of PDF sets and their errors

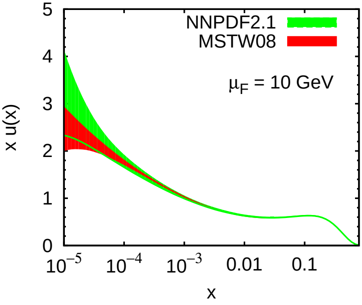


error bands for 68% CL

- spread between different parameterizations often larger than error bands of single parameterization
- error bands propagate uncertainties of fitted data into PDFs but do not reflect "systematic theory uncertainties" of PDF extraction

Illustration of PDF sets and their errors

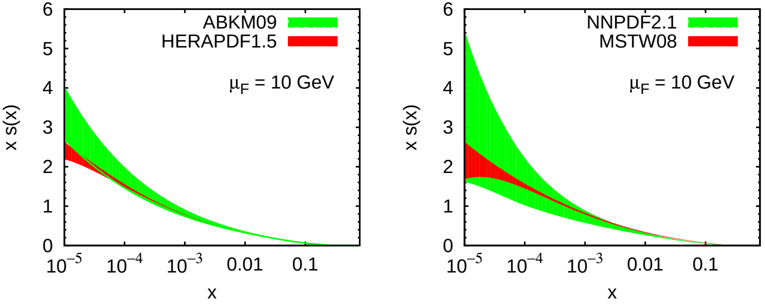




error bands for 68% CL

- spread between different parameterizations often larger than error bands of single parameterization
- error bands propagate uncertainties of fitted data into PDFs but do not reflect "systematic theory uncertainties" of PDF extraction

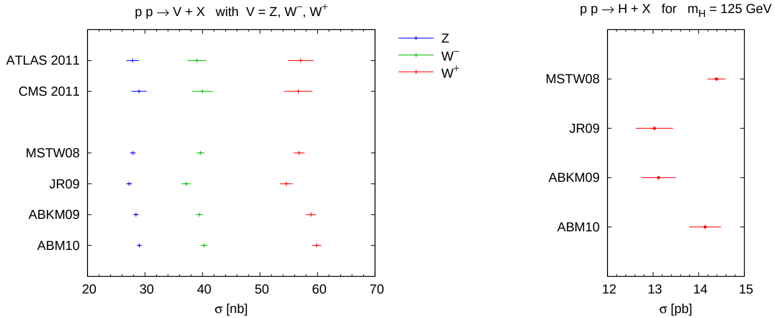
Illustration of PDF sets and their errors



error bands for 68% CL

- spread between different parameterizations often larger than error bands of single parameterization
- error bands propagate uncertainties of fitted data into PDFs but do not reflect "systematic theory uncertainties" of PDF extraction

Illustration of PDF sets and their errors



benchmark NNLO cross sections at $\sqrt{s} = 7\,\mathrm{TeV}$ from Alekhin et al., arXiv:1011.6259

- spread between different parameterizations often larger than error bands of single parameterization
- error bands propagate uncertainties of fitted data into PDFs but do not reflect "systematic theory uncertainties" of PDF extraction

Summary

- estimating theoretical uncertainties $\neq$ an exact science
- "scale uncertainty" based on renormalization group eq.
  estimates certain higher-order terms in α_s
  prescriptions for scale choice = educated guesses
- higher orders in pert. theory not the only source of uncertainty
  full final state details, hadronization corrections, ...
  are more difficult to quantify
- errors of PDF fits reflect uncertainties of fitted data
  (not always a straightforward exercise in textbook statistics)
  do not include uncertainties of theory used to fit data