QCD aspects of hadron collider physics

Kirill Melnikov

Johns Hopkins University

APS meeting, Denver, May 3, 2009

Preface

$$\mathcal{L}_{\mathrm{QCD}} = -\frac{1}{4} G^a_{\mu\nu} G^{a,\mu\nu} + \sum_{i=1}^{n_f} \bar{\psi}_i \left(i\hat{D} - m_i \right) \psi_i$$

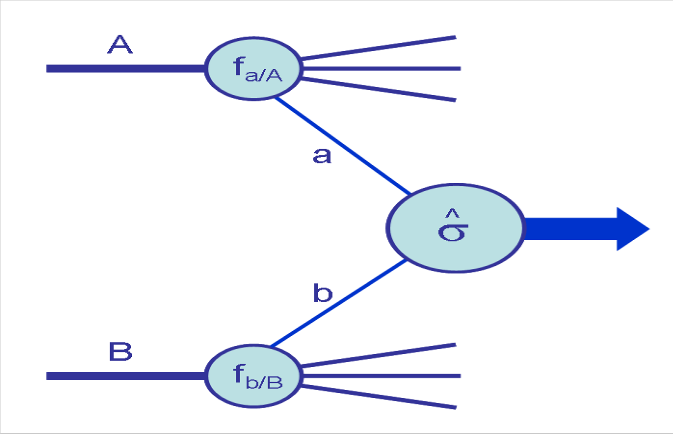
- QCD is a rich theory that describes many phenomena around us
- It is the only non-perturbative QFT that we can access in the laboratory
- Testable QCD predictions span distances from 1 fm to 0.001 fm and energies from few hundred MeV to few TeV
- QCD will be responsible for much of the LHC physics
- This talk – QCD at the Tevatron and the LHC

Preface

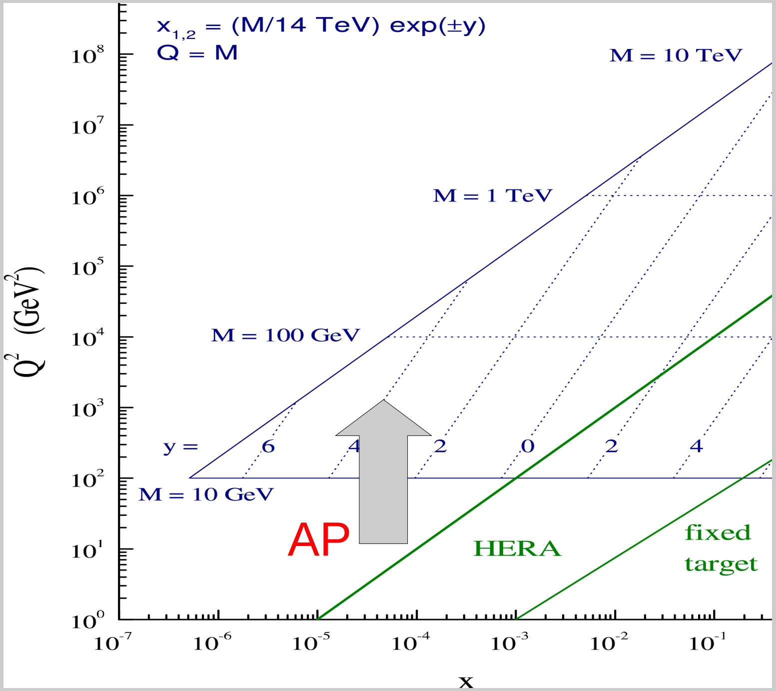
- Many hadron collider QCD applications are based on the factorization theorem

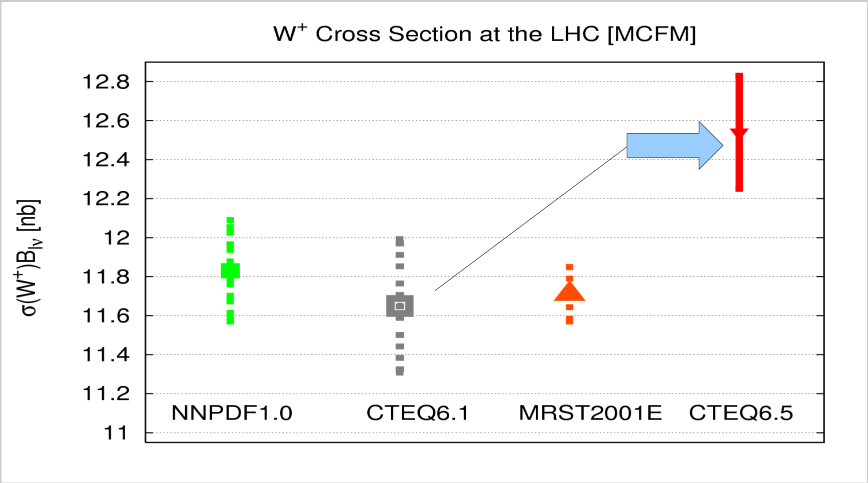
$$\langle \mathcal{O} \rangle = \sum_{i,j} \int \mathrm{d}x_1 \mathrm{d}x_2 \; f_i(x_1) f_j(x_2) \; \mathrm{d}\sigma_{ij \to p} \; \mathcal{F}_{p \to \mathcal{O}} + \mathcal{O}(1/Q)$$

- According to the theorem, we have the following objects to deal with
  - parton distributions
  - hard scattering cross-sections
  - fragmentation (parton showers, hadronization, jets)
  - higher twist, underlying event (beyond the unproven theorem)



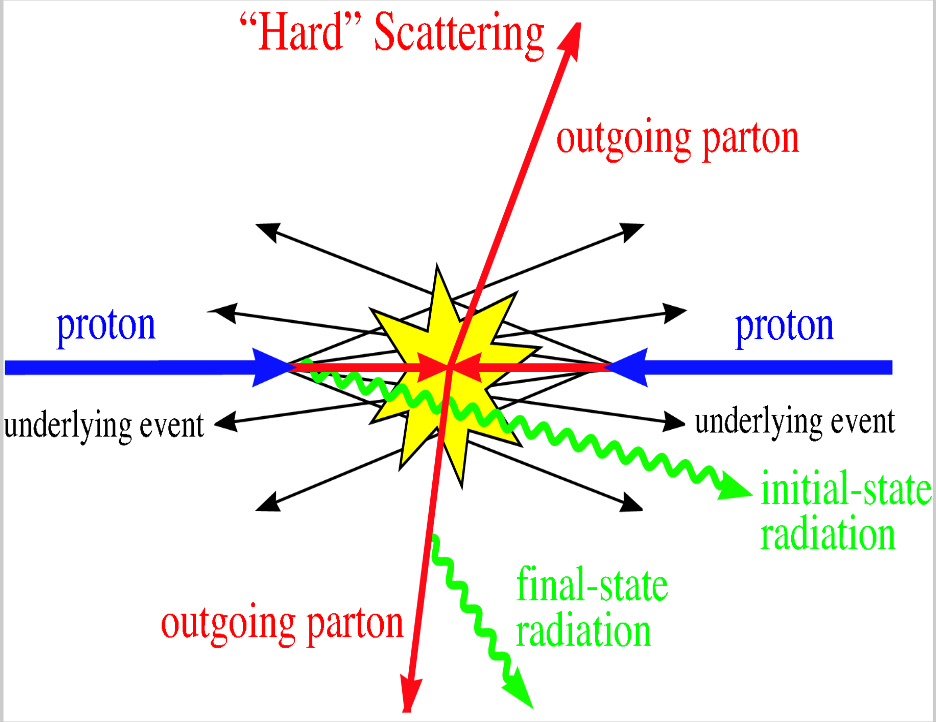
Parton distributions





- Parton distributions are non-perturbative objects determined from dedicated fits to data
- MTSW, CTEQ, Alekhin, NNPDF
- Issues:
  - consistency of data sets
  - Initial parameterization bias
  - PDF error estimates
- Examples:
  - MTSW 2008 update reduced Tevatron Higgs production cross-section by about 10 %
  - CTEQ6M increased the W production cross-section at the LHC by about 6 %

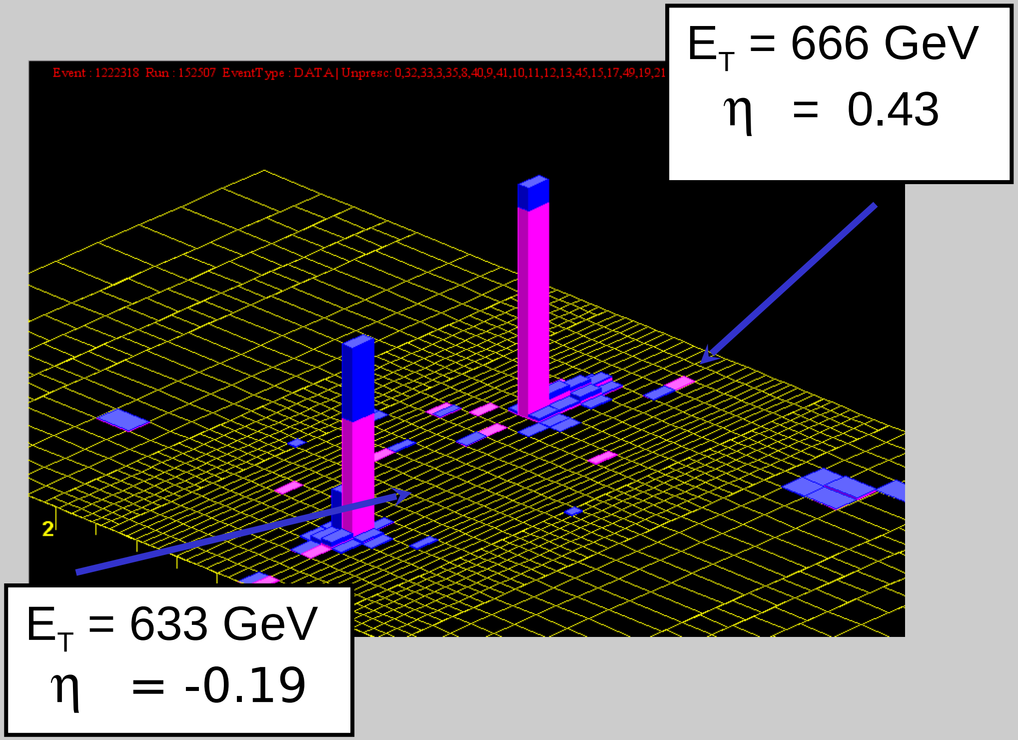
Hard scattering



$$d\sigma = d\sigma_0 + \frac{\alpha_s}{\pi}\left(d\sigma_{12}\ln^2\frac{p}{\lambda} + d\sigma_{11}\ln\frac{p}{\lambda} + d\sigma_{10}\right) + +\mathcal{O}(\alpha_s^2)$$

- Thanks to asymptotic freedom, hard scattering can be computed in perturbation theory
- Different approximations – parton showers, fixed (LO, NLO, NNLO) order matrix elements and combinations of the above
- What is applicable where? What is the accuracy?
- Recent highlights
  - consistent framework for combining parton showers and fixed order computations
  - breakthrough in NLO computations

Fragmentation



One of the di-jet events with the highest jet transverse momenta

- Change of degrees of freedoms: quarks and gluons → hadrons
- Jets, hadronization, underlying event
- Description relies on parton showers, resummations, models
- Study of underlying event with the Tevatron data
- Infra-red unsafe algorithms are being phased out; fast implementations of seedless cone-based infra-red safe algorithms are available
- Focus on jet algorithms designed to achieve specific physics goals (study QCD, discover new particles, etc.)

Preface

- QCD for hadron colliders is a vibrant and broad field
- Since it is impossible to do justice to all topics, I decided to discuss topics where I believe real breakthrough occurred in the past year
  - next-to-leading (NLO) computations in QCD
  - QCD ideas for Higgs and BSM discoveries at the LHC
  - new approaches to parton distribution functions determination

NLO computations

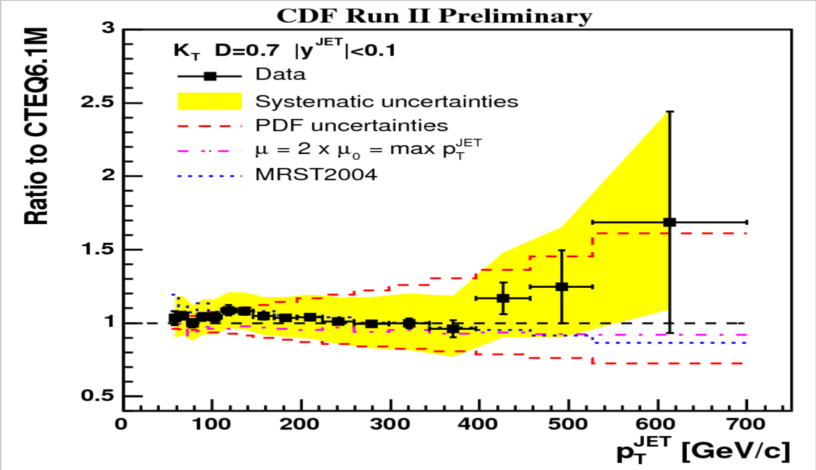
- There are many ways to compute hard scattering cross-sections and NLO is just one of them
  - parton showers (PYTHIA, HERWIG, SHERPA)
  - resummations (RESBOS, etc)
  - fixed order computations (ALPGEN, MADGRAPH, HELAC, COMIX, MCFM)
  - … and combinations of the above (CKKW, MC@NLO, POWHEG)
- These approaches are based on different approximation and have different regions of applicability
- Roughly:
  - Parton showers & resummations → phase space edges (soft, collinear)
  - Fixed order calculations (LO, NLO) → bulk of the phase space
- A proper tool to describe an observable depends on the observable

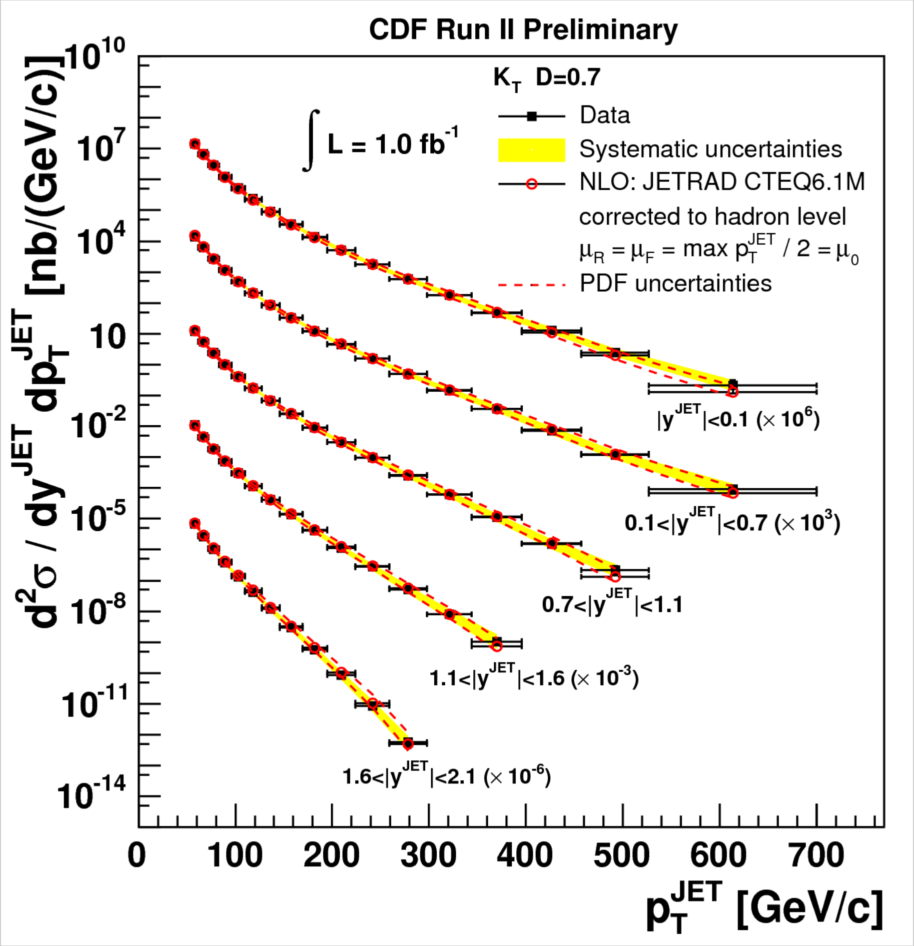
NLO computations

- Many observables relevant for BSM searches at the Tevatron and the LHC are dominated by large momentum transfers → NLO QCD should be the right tool
- The problem is that we don't know what ``large momentum transfer'' really means but we can learn this from the Tevatron data
- CDF and D0 started releasing comparisons between theory and experiment based on ~1 fb^{-1} of integrated luminosity. Adanced theory is employed:
  - ALPGEN or MADGRAPH matched to PYTHIA or HERWIG are used for leading order studies
  - MCFM, MC@NLO, JetRAD, NLOJET++ and other dedicated NLO routines are used for NLO predictions
- The goal of these studies is to establish what works and what doesn't and to draw some lessons for the LHC

NLO: inclusive jet production

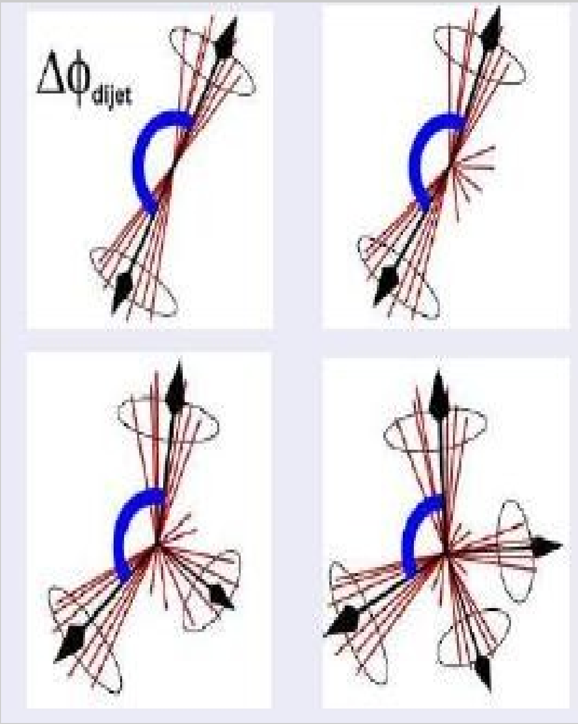
- Measurement of inclusive jet transverse energy distribution is a classic result sensitive to the interplay of
  - NLO computations
  - large-x PDFs
  - BSM four-fermion contact interactions
- D0 & CDF measurements are in good agreement with NLO QCD

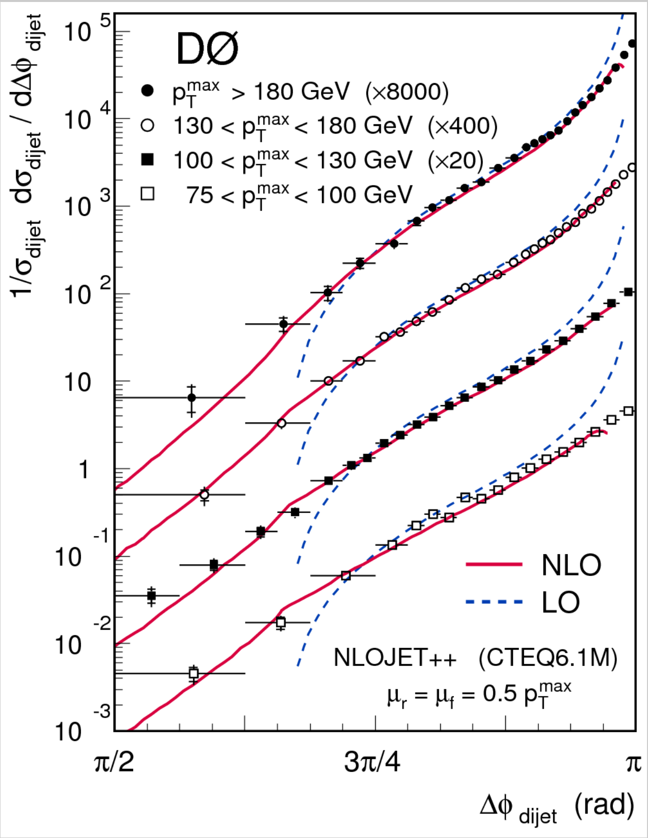


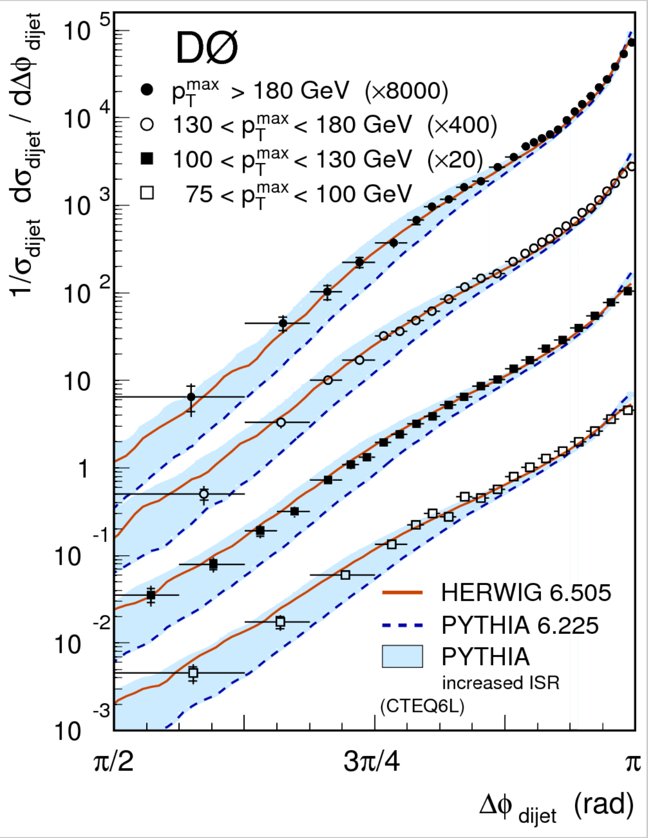


NLO: dijet azimuthal correlations

- $\Delta\varphi$ distribution is sensitive to multiple soft emissions at $\Delta\varphi=\pi$ and to hard emissions for smaller $\Delta\varphi$
- A single plot allows us to check various pQCD approximations



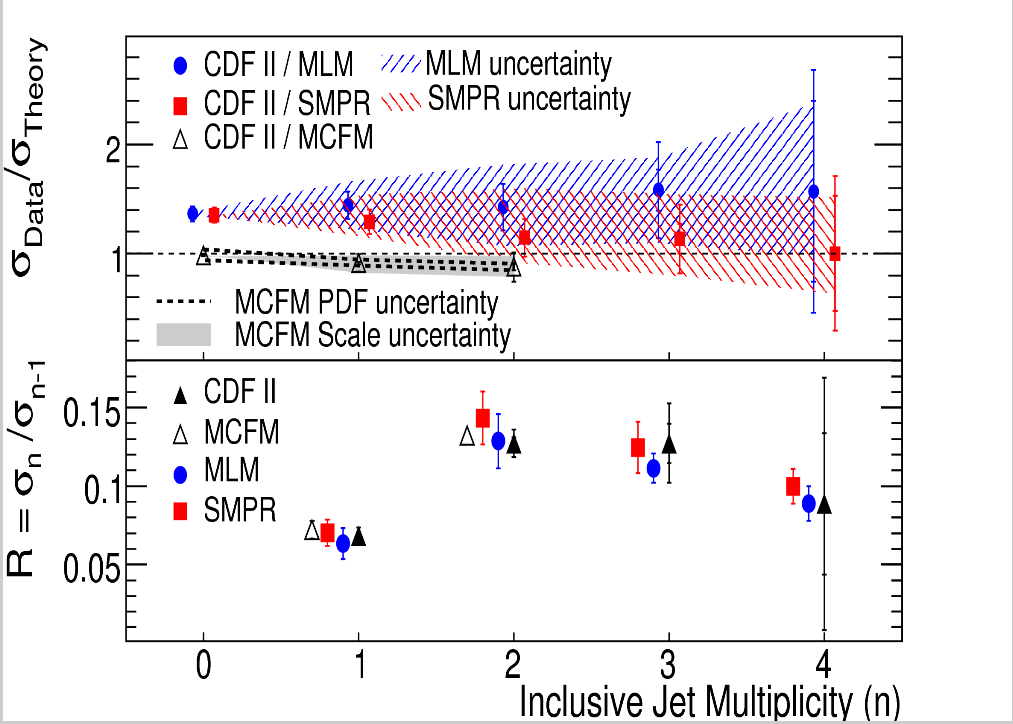




MC@NLO, POWHEG

NLO QCD: W+jets at the Tevatron

- W+jets is another interesting example of where we stand
- It is an important background for many Tevatron and the LHC searches
- CDF data agrees very well with LO computations matched to parton showers and exceptionally well with NLO QCD
- NLO QCD results exhibit very small scale dependences



NLO computations

- NLO QCD predictions for W+3 and W+4 jets used to be beyond computational capabilities
- NLO QCD computations with large number of particles are simple, as a matter of principle but hard, as a matter of practice
- Whether or not a particular process can be computed through NLO QCD is a function of multiplicity; there is a sharp cut-off at $2 \to 4$ processes:
  - many $2 \to 3$ processes computed (3jet, Hjj, WWj, VVV, ttZ, etc.)
  - .. but not a single $2 \to 4$ process is fully known through NLO QCD
- Several reasons, that we keep citing for many years
  - many diagrams
  - complicated analytic treatment
  - numerical instabilities
- But, there is a feeling that the situation started to change rapidly

NLO QCD computations

- During the past year computational methods became sufficiently mature to take on 2 → 4 processes
  - pp → tt bb — Denner, Dittmaier, Pozzorini
  - pp → W +3 jets — Blackhat (Bern et al.), Rocket (Ellis et al.)
- These are very important results – they open the door for NLO QCD computations to many 2 → 4 processes and beyond
- Traditional Passarino-Veltman reduction-based methods were developed and optimized over the past 15 years; they are responsible for the bulk of existing NLO phenomenology
- Unitarity is a new game in town but it is very promising and is also interesting since it changes perspective on how QFT works [at one-loop]

NLO QCD computations

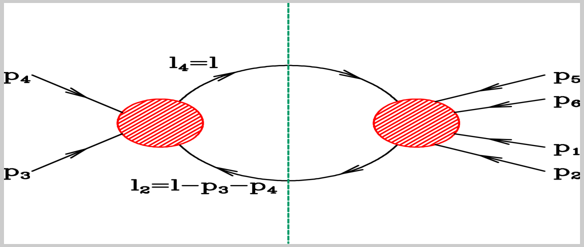
- Computation of one-loop corrections is the bottleneck
- NLO computations seek to determine reduction coefficients c_j

$$\mathcal{A}^{1-\mathrm{loop}} = \sum c_j I_j$$

- Unitarity is helpful because
  - It constrains reduction coefficients
  - tree-level amplitudes are involved in the constraint

$$\mathcal{A}^{1-\mathrm{loop}} = \sum c_j I_j \qquad \mathrm{Im}\left(\mathcal{A}^{1-\mathrm{loop}}\right) \propto \sum \left|\mathcal{A}^{\mathrm{tree}}\right|^2$$

$$\sum c_j \mathrm{Im}(I_j) \propto \sum \left|\mathcal{A}^{\mathrm{tree}}\right|^2$$



  - In the past few years, it was understood how to use such constraints efficiently

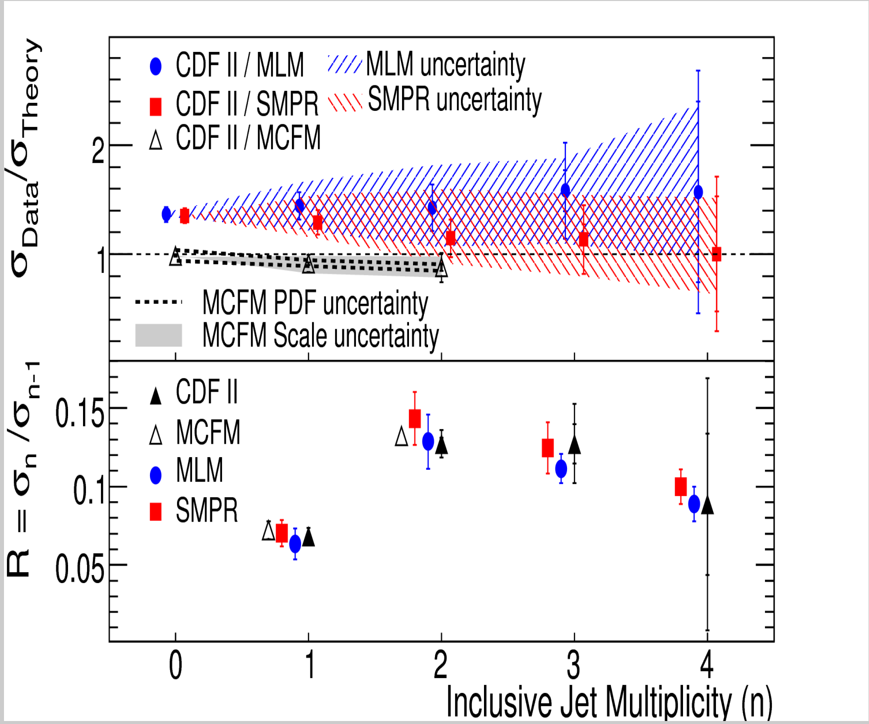
NLO computations

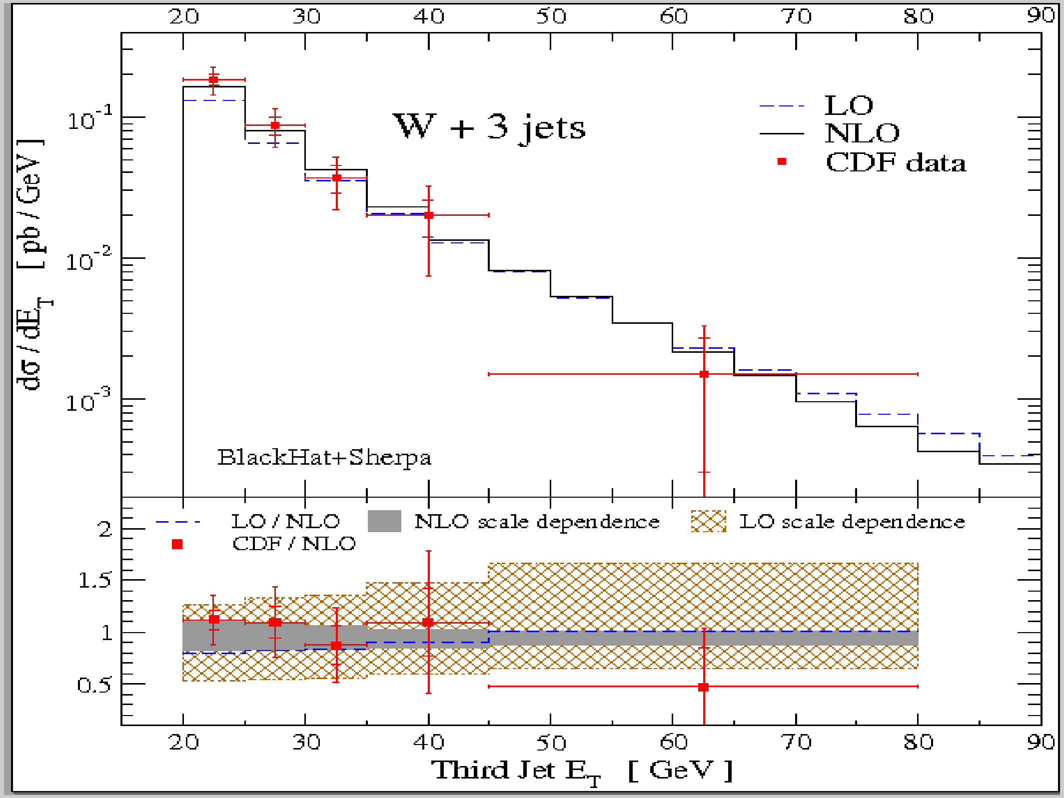
- Bern, Dixon and Kosower envisioned importance of unitarity (~ 1990)
- Impressive early successes (5 partons, Z(W)+4 partons) but no real computational method. Impact limited
- In the past few years several observations resulted in a breakthrough
  - Cachazo, Britto, Feng point out that one can take cuts of one-loop amplitude with respect to loop momentum (rather than external kinematic invariants)
  - Ossola, Pittau and Papadopoulos (OPP) showed that reduction coefficients c_j can be reconstructed if integrands of Feynman integrals are known for special values of the loop momenta, where at least one inverse Feynman propagator vanishes ↔ at least one on-shell particle
  - Ellis, Giele and Kunszt showed that OPP reduction procedure meshes well with generalized unitarity
- These and other developments resulted in generalized unitarity becoming a practical tool for phenomenology

W+3 jets at the Tevatron

- First physics applications of unitarity methods – W+3 jet production at the Tevatron and the LHC (Rocket, Blackhat)
- Excellent agreement with the CDF data

Bern, Berger, Dixon, Kosower, Forde, Febres-Cordero, Ita, Maitre, Gleisberg





Implications for the LHC

- NLO QCD successes in describing Tevatron data should encourage us to trust NLO QCD predictions for the LHC
- Unitarity techniques enable NLO QCD computations for high-multiplicity processes, enhancing degree of realism that NLO can provide

An experimenter's wishlist

- Hadron collider cross-sections one would like to know at NLO

Run II Monte Carlo Workshop, April 2001

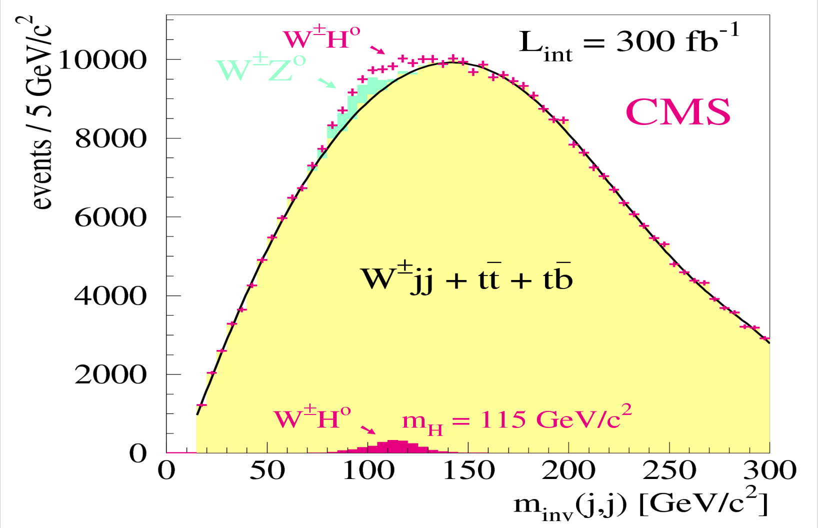
Single boson	Diboson	Triboson	Heavy flavour
$W + \leq 5j$	$WW + \leq 5j$	$WWW + \leq 3j$	$t\bar{t} + \leq 3j$
$W + b\bar{b} + \leq 3j$	$WW + b\bar{b} + \leq 3j$	$WWW + b\bar{b} + \leq 3j$	$t\bar{t} + \gamma + \leq 2j$
$W + c\bar{c} + \leq 3j$	$WW + c\bar{c} + \leq 3j$	$WWW + \gamma\gamma + \leq 3j$	$t\bar{t} + W + \leq 2j$
$Z + \leq 5j$	$ZZ + \leq 5j$	$Z\gamma\gamma + \leq 3j$	$t\bar{t} + Z + \leq 2j$
$Z + b\bar{b} + \leq 3j$	$ZZ + b\bar{b} + \leq 3j$	$WZZ + \leq 3j$	$t\bar{t} + H + \leq 2j$
$Z + c\bar{c} + \leq 3j$	$ZZ + c\bar{c} + \leq 3j$	$ZZZ + \leq 3j$	$t\bar{b} + \leq 2j$
$\gamma + \leq 5j$	$\gamma\gamma + \leq 5j$		$b\bar{b} + \leq 3j$
$\gamma + b\bar{b} + \leq 3j$	$\gamma\gamma + b\bar{b} + \leq 3j$		
$\gamma + c\bar{c} + \leq 3j$	$\gamma\gamma + c\bar{c} + \leq 3j$		
	$WZ + \leq 5j$		
	$WZ + b\bar{b} + \leq 3j$		
	$WZ + c\bar{c} + \leq 3j$		
	$W\gamma + \leq 3j$		
	$Z\gamma + \leq 3j$		

QCD ideas and BSM searches

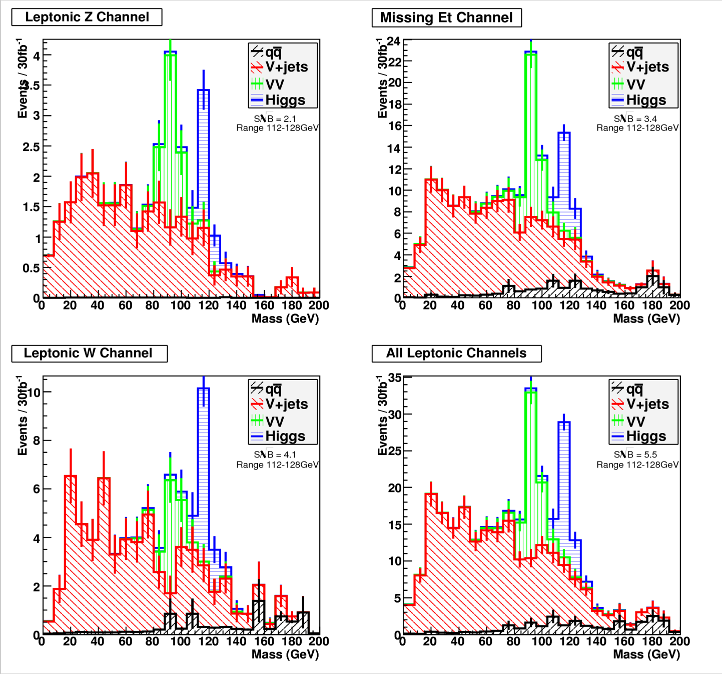
- Understanding of QCD should enable us to broaden BSM search strategies
- A lot of recent work in that direction
  - appearance of new discovery channels (Higgs, resonances)
  - design of new jet algorithms
- An example: can one discover a Higgs boson through its decay to bottom pairs at the LHC?
  - Conventional wisdom says – no way since it will be swamped by direct bottom production
  - But, ingenious suggestion by Butterworth, Davidson, Rubin and Salam opens up a window of opportunity

PP → WH & H → bb

- Recall why searching for pp → WH(bb) is hard
  - σ(pp → WH(bb)) ≈ few pb , σ(pp → W jj) ≈ few x 10^4 pb
  - σ(pp → Wbb) ≈ few pb σ(pp → tt) ≈ 800 pb, σ(pp → bt) ≈ 400 pb
- Signal extraction is clearly very difficult. The question is can one do significantly better. The right panel shows that it is possible



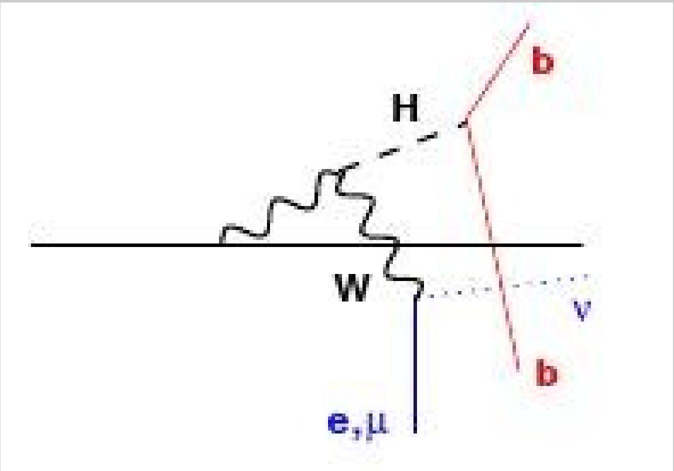
ATLAS TDR

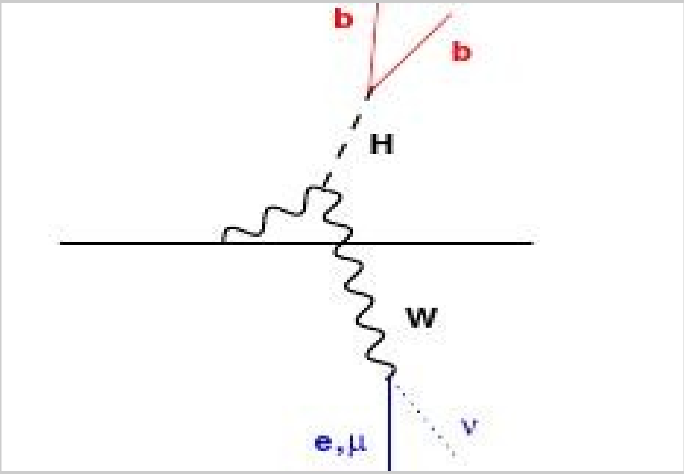


Butterworth, Davidson, Rubin, Salam

PP → WH and H → bb

- Here are three main ideas
  - require high-p_T W boson and the Higgs boson in the event
    - leads to back-to-back events where two b-quarks are contained within the same jet .
    - high-p_T does decrease the signal BUT it reduces the background even stronger (e.g. kills tt production)
    - improves acceptancies and kinematic resolution

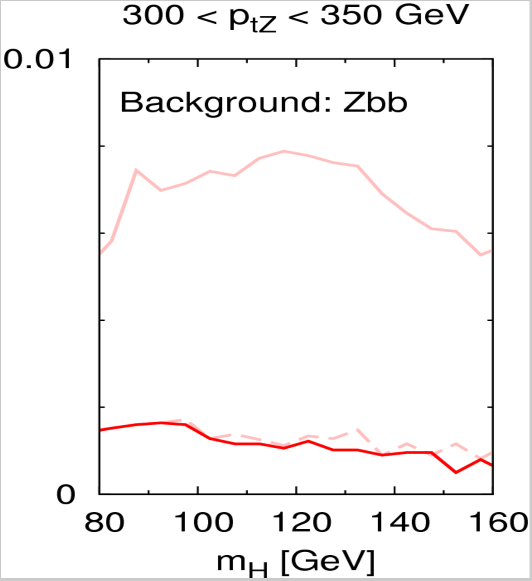




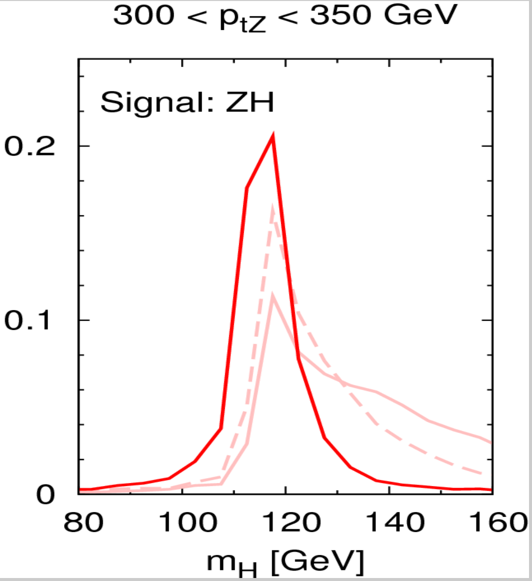
PP → WH and H → bb

- use differences in branching patterns of H → bb and g → gg, q → qg,..
  - QCD partons prefer soft emissions:
    parent → hard off-spring + soft off-spring
  - H → bb prefers (energy) symmetric branchings
- use jet algorithm which reflects underlying pattern of QCD radiation to improve the mass resolution
  - beat down contamination from the underlying event
  - capture most of perturbaive QCD radiation

Mass drop

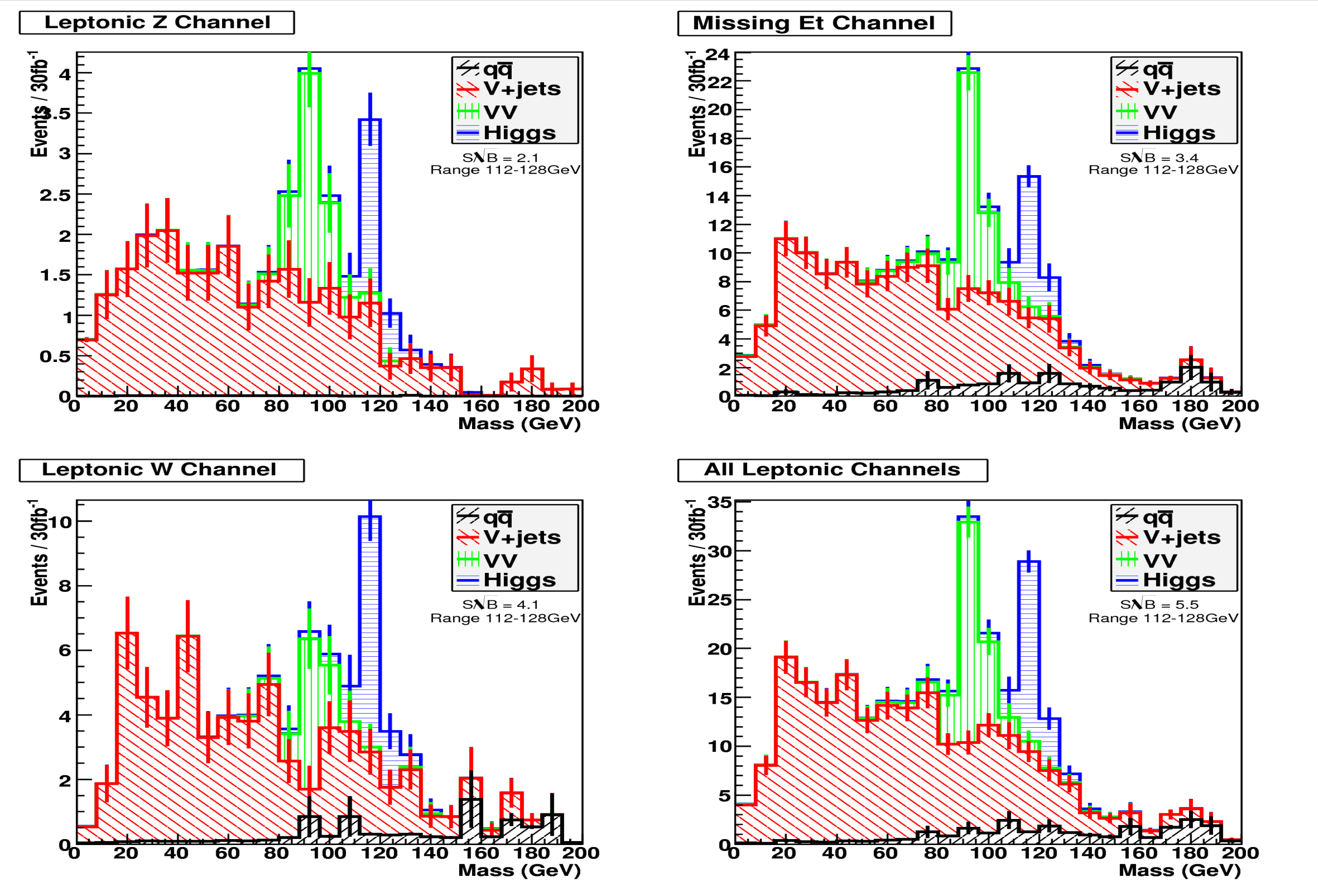


filtering



PP → WH and H → bb

- And, when this method is employed, the results are spectacular

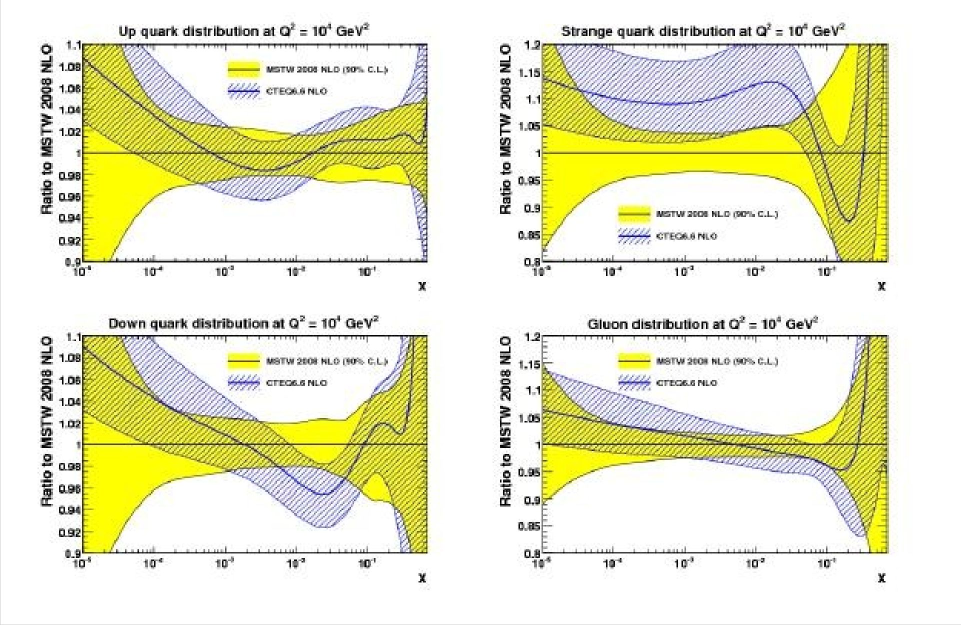


Techniques are generic; related studies by other groups

Thaler and Wang, Kaplan, Rehermann, Scwartz, Tweedie, Ellis, Vermilion, Walsh, Almedia, Perez, Lee, Sterman, Virzi, Sung

Parton distribution functions

- Parton distribution functions (PDFs) are important for any predictions related to hadron collider physics
- Experimental precision and theoretical calculations improved to a degree that good understanding of PDFs and their errors is relevant
- PDF fitting became an industry with such ``brand'' names as MRST-MSTW, CTEQ and Alekhin competing for global customers
- Often PDFs by different groups are incompatible even within errors



Parton distribution functions

- This is somewhat surprising since all ``brand names'' use similar strategies in PDF fitting
  - choose initial parameterization $f(x) \propto x^{\alpha}(1-x)^{\beta}$
  - choose data set to be fitted (DIS, Drell-Yan, Z, W, Tevatron jets)
  - determine parameters of the parameterization and their uncertainties
- Incompatible results occur because of subjective choices in the above procedure. Different groups
  - include different data;
  - employ different initial parameterizations;
  - use different criteria for PDF uncertainty estimates
- More objectivity in dealing with these issues is desired

Parton distribution funtions

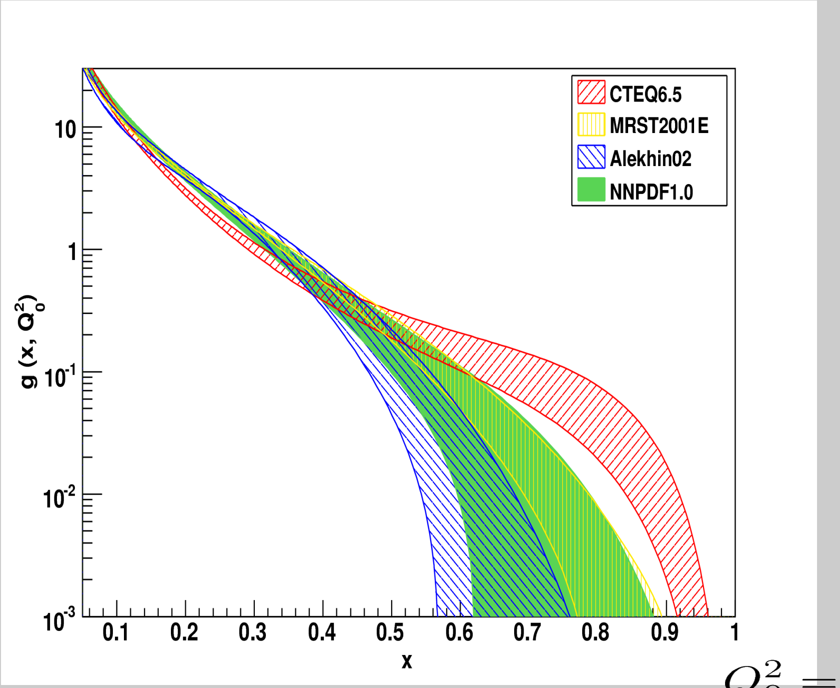
- One option is to introduce probability distributions for PDFs and to infer these probability distributions from available data Giele, Keller, Kosower
- If the probability distributions are known, it is easy to generate statistical ansamble of PDFs and to compute PDF-related uncertainties for a any observable
- If the probability distribution properly samples functional space of PDFs, initial parameterization bias is completely avoided
- This procedure requires mapping of infinitely-dimensional space by finite number of measurements. So the question is how to implement this in practice

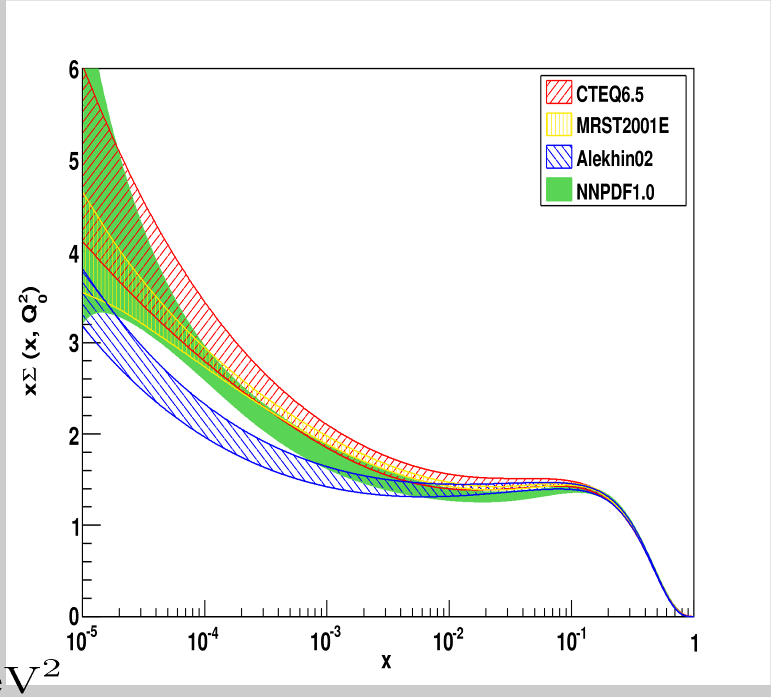
Parton distribution functions

- NNPDF collaboration suggested to use a combination of Monte-Carlo techniques and neural network for efficient sampling and initial parameterization de-bias
  - existing data is used to generate artificial data sets large enough so that actual data and errors are obtained using standard averages
  - artificial data sets are employed to construct N_s parton distribution functions using neural network output
  - To this end, data is divided into training and validation samples and the procedure stops when fit quality in validation sample does not increase
  - A set of N_s PDFs is used to calculate observables and their PDF-related uncertainties

Parton distribution functions

- This procedure leads to interesting features
  - reasonable agreement with brand names central values
  - compared to brand name fits, errors are typically large; in particularly for such values of x where data is absent
  - procedure handles incompatible data sets by terminating fitting when no improvement in validation set can be achieved
  - stable against ``data removal'' – error increases, central value stays





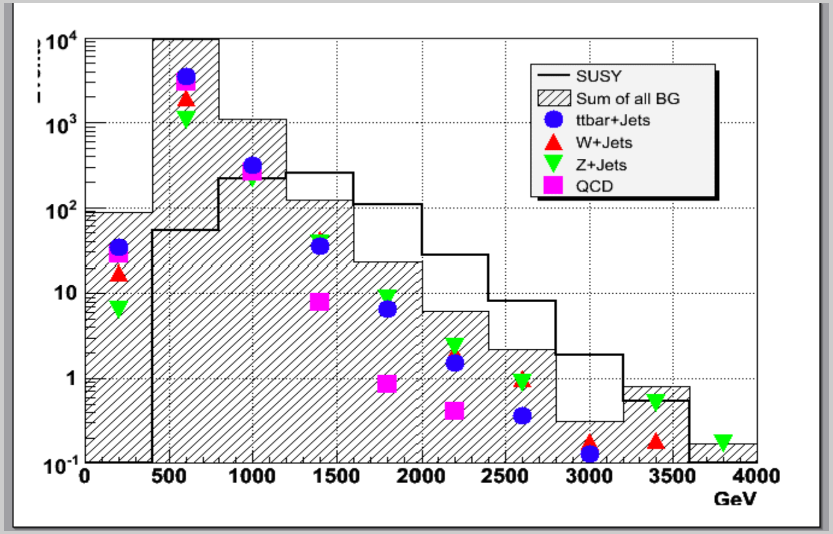
$$Q_0^2 = 2\ \mathrm{GeV}^2$$

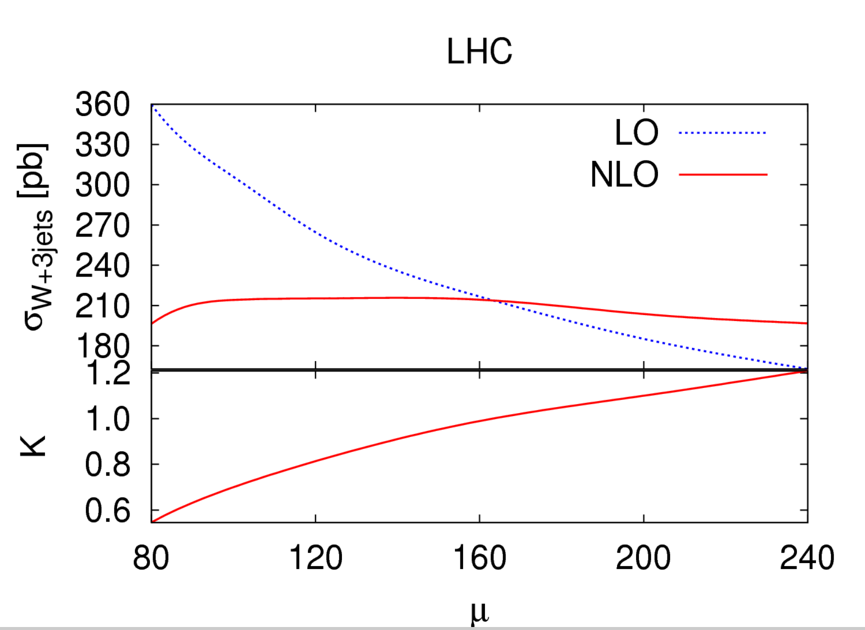
Conclusions

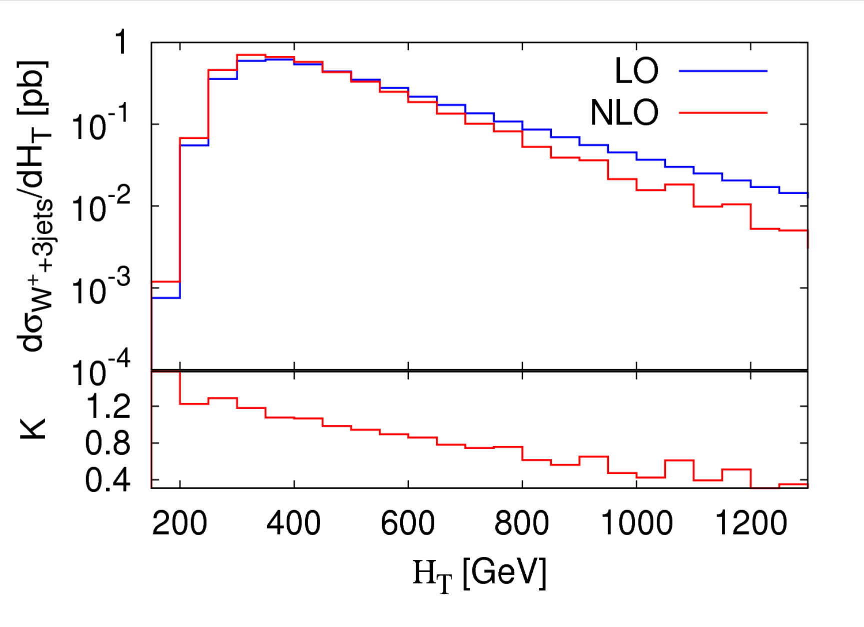
- There is little doubt that QCD will play very important role at the LHC
- QCD can describe Tevatron data very well if theoretical tools are used as appropriate. This gives encouragement for the LHC
- Last year brought remarkable progress
  - very important developments in NLO computations – new computational methods were finally put to work and $2 \rightarrow 4$ processes are within reach
  - interesting developments in designing jet algorithms to satisfy particular physics goals
  - new ideas related to PDFs determination from data
  - ..and many other things I did not talk about (NNLO, MC+NLO, CKKW, studies of underlying event, etc)

W+3 jets at the LHC

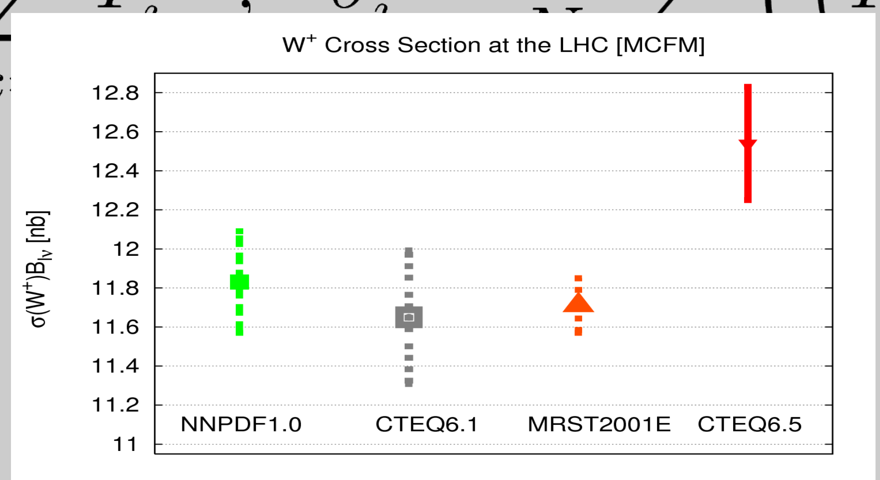
- Comparison of NLO QCD predictions for W+3 jets with the Tevatron data should encourage us to trust NLO QCD predictions for the LHC







$$F_i = \frac{1}{N_s}\sum_{k=1}^{N_s} F_i^{(k)}, \quad \sigma_i^2 = \frac{1}{N_s}\sum^{N_s}\left(\left(F_i^{(k)}\right)^2 - F_i^2\right).$$



NLO computations

- F90 program Rocket can compute the following one-loop amplitudes

  Ellis,Giele, Kunszt, Melnikov, Zanderighi

  - N-gluon scattering amplitudes
  - two-quark (massless and massive) + N-gluon scattering amplitudes
  - W+two-quark+N-gluon amplitudes
  - W+four-quarks+1 gluon
  - massive quark pair + massless quark pair + N gluons
- Note that N is a PARAMETER which is specified alongside with collision energy, polarization states, masses etc.
- To have the right perspective: for all-gluon amplitudes look at the number of Feynman diagrams
  - N=6 → 10860 Feynman diagrams
  - N=7 → 168 925 Feynman diagrams
  - Rocket was successfully used to compute an N=20 gluon amplitude!

Giele, Zanderighi