First Year Transfer Report: Next-to-Leading-Order and Monte-Carlo Predictions for Inclusive Cross-Sections at the LHC

Oliver Harris

June 4, 2004

1 Abstract

The purpose of this report is to outline the progress I have made in the first eight months of my postgraduate research degree, and also the directions in which I plan to continue my investigations. I present inclusive dijet production cross-sections in *pp* collisions at LHC energies, calculated using Frixione's and Ridolfi's FORTRAN jet production package (at next-to-leading order, or NLO), and plans for work with the pQCD calculation/Monte Carlo 'hybrid' program, MC@NLO, combined with Atlfast (The ATLAS fast simulation).

Contents

2 Introduction

When CERN's Large Hadron Collider (LHC) goes online in 2007, it will be the world's largest and most powerful particle accelerator, colliding protons with a centre-of-mass energy of 14 TeV. Protons, being hadrons consisting of quarks and gluons, can interact with each other in complicated ways, and one of the most important things to understand in computing hadronic cross-sections is the particles' PDF. This is a function, derived partly from theoretical predictions but mostly from experimental data, that describes how much of a hadron's momentum is carried by its various partons (elementary constituents). There are three classes of partons:

i) valence quarks, the three 'real' quarks that make up the proton, two u flavour and one d flavour;

ii) the quark 'sea', virtual quark-antiquark pairs that condense out of the gluonic field binding the three valence quarks. These come in all flavours, not just the lightest two. Although the heavier flavours are strongly suppressed at low energies, contributions from charm and bottom quarks become significant in hard scatters at energies in the LHC regime. Top contributions can safely be ignored;

iii) gluons, the gauge bosons of QCD that mediate the strong force.

PDFs relate a parameter called x (the fraction of the incoming hadron's momentum carried by each type of parton, in the 'infinite momentum' frame of the hadron) as a function of energy in the collision's centre-of-mass frame. [1] It is additionally complicated by the fact that x varies not only with collision energy but also with the parameter Q^2, which represents (the square of) the momentum exchanged in the hard scatter, whether this is an s-, t- or u-channel event. (In lepton-hadron collisions, which occur at HERA, Q^2 represents the momentum carried by the exchanged Z^0, $W^{\pm}$ or photon; in hadron-hadron collisions, dozens of different processes are possible, so there is no unique or preferred definition of Q^2.)

A PDF, at a given value of Q^2, can be graphically represented as a quantity called the form factor $f_p(x)$ (the probability that a parton of flavour p will have a particular value of x) plotted against x, with different lines representing valence quarks, sea quarks and gluons. In practice, the y-axis variable is usually plotted as $xf_p(x)$, so that it represents the 'amplitude-weighted' momentum carried by each type of parton. Naturally, PDFs are normalised (in that the integrated momentum contributions from all parton types sum to unity). An example of a graphed PDF is given below:

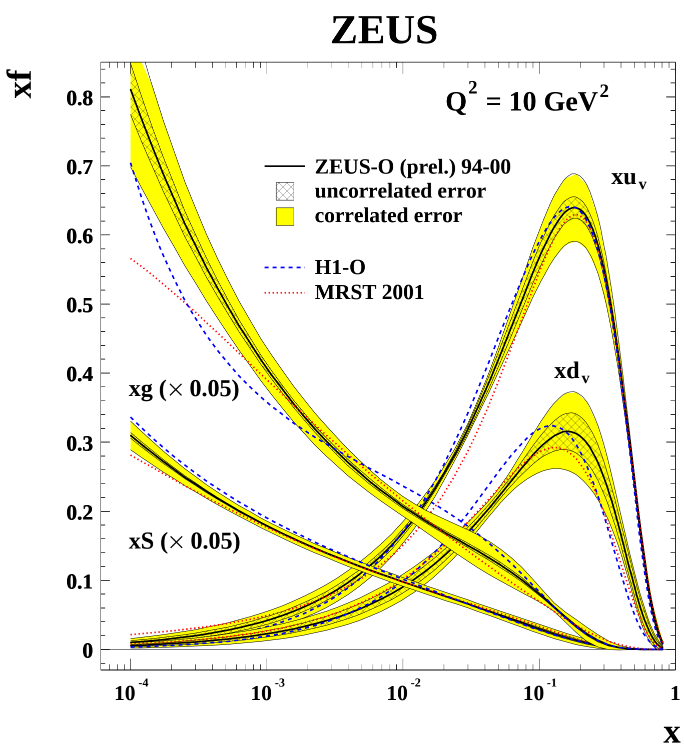

Figure 1: A PDF based on data taken with ZEUS, at $Q^2 = 10$ GeV^2; partons are labelled d_v (valence, down), u_v (valence, up), S (sea, all flavours) and g (gluons). The PDF sets H1-O and MRST 2001 are plotted for comparison.

Only at high Q^2 do contributions from heavier quark flavours (c and b) become significant. There are, in general, two types of PDF scheme, classified according to the type of pQCD momentum renormalisation scheme involved in the theoretical calculations: DIS (Deep Inelastic Scattering) and $\overline{MS}$ (modified Minimum Subtraction). The distinction need not concern us here.

There are many PDF schemes currently available, identified by names which are based on the authors of the scheme and some sort of version number. Two of the most important groups are MRST (Martin, Roberts, Stirling and Thorne) and CTEQ (Coordinated Theoretical-Experimental project on QCD), which consists of some two-dozen world experts on QCD. The different schemes vary not only in the way they evaluate $F_p(x)$ at different values of x and Q^2, but in a number of other parameters, such as n_f, the number of quark flavours considered; Λ_{QCD}, the characteristic QCD scale factor; α_s, the strong coupling constant; and the order of α_s to which calculations are taken (i.e. LO or NLO).

2.1 The importance of PDFs in HEP experiments

An accurate, precise PDF is extremely important in particle physics experiments, especially those at very high energies in which entirely new types of physics (Higgs bosons, supersymmetry, large extra dimensions, leptoquarks etc) are expected to appear. Specifically, errors on f at high-x must be kept as low as possible, as it is at these values that new physics will occur (as the energy is greater in the parton CoM frame). If a model predicts, for some signature process, a given deviation in, say, cross-section (from the standard model) at a given parton energy (that is, a given value of x for incoming hadrons with fixed energy), this will only be discoverable if the deviation is larger than that due to uncertainty in the PDF. As many predicted physics processes have associated observables whose magnitude grows monotonically with parton CoM energy, a more precise PDF essentially lowers the threshold parton energy at which the process may be discovered. I plan to use PDFs based on data taken with the ZEUS experiment at HERA to constrain discovery potentials for various new physics models, such as those mentioned above.

This is where measurements taken with the experiments on HERA (*H*adron-*E*lektron *R*ing *A*nlage) at DESY, Hamburg, have an impact on physics searches at the LHC; without accurate and precise PDFs, LHC will be a powerful but blunt tool. Data taken at HERA will transform LHC from a sledgehammer into a scalpel.

3 Athena and Atlfast

The ATLAS experiment (*A* *T*oroidal *L*HC *A*pparatu*s*) is one of four experiments being constructed on the Large Hadron Collider. It is a large, state-of-the art multipurpose detector, with four layers of components: the inner tracker, electromagnetic and hadronic calorimeters, and muon detector.

Athena is the object-oriented control framework, based on the GAUDI architecture originally developed for the LHCb experiment, in which software for ATLAS is developed. The purpose of this framework is to provide a consistent, stable environment for users to write their code (event simulations, detector simulations, visualisation programs etc.) which is simultaneously easy-to-use, flexible and easily updatable as new technology becomes available.

Central to the Athena framework is CMT:

> "CMT is a configuration management environment, based on some

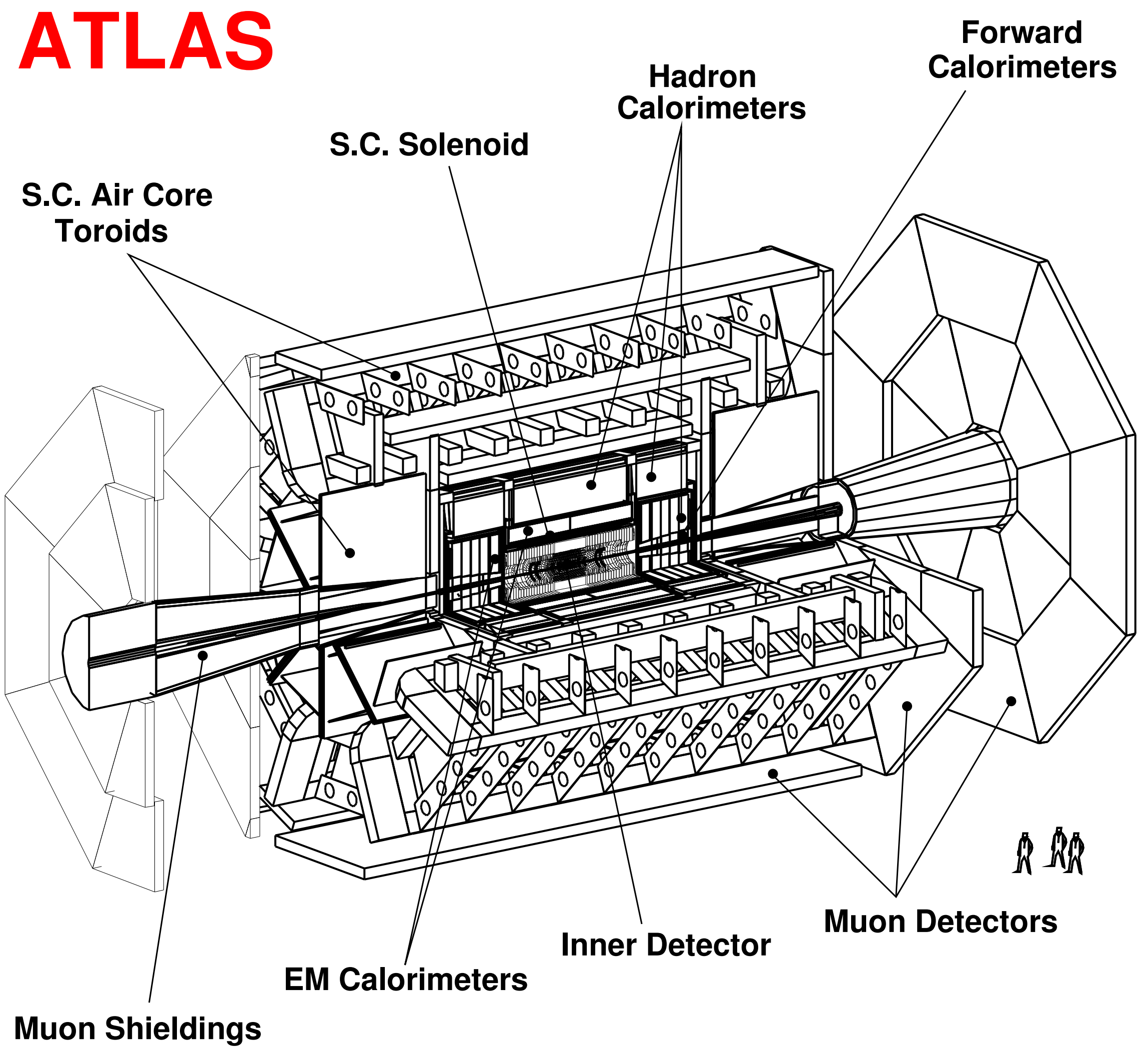

Figure 2: The ATLAS experiment

> management conventions and comprises several shell-based utilities. It is an attempt to formalise software production and especially configuration management around a package-oriented principle." [2]

Atlfast is the ATLAS fast simulation. When run, the program calls a Monte Carlo (typically one of the standard generators - Pythia, HERWIG etc) to produce 'truth' events, then runs this data through a simulation of ATLAS. Truth data, in this sense, is data as it would be picked up by a hypothetical perfect detector - that is, partons have been allowed to hadronise and resonances allowed to decay, but nothing else has been done to the data.

4 Progress made so far

My initial investigations have so far proceeded down two routes; firstly, the Athena environment and the Atlfast simulation, and secondly the NLO package, which calculates cross-sections for both leading-order and next-to-leading order QCD events (typically, NLO events are the same as LO events but with an extra gluon emitted in either the initial or final state; the overall amplitude for the relevant diagram therefore has an extra factor of α_s, hence the term 'next-to-leading-order). Eventually, I hope to do a full Monte Carlo analysis with MC@NLO: this is the first package to combine Monte Carlo with NLO calculations, making use of the HERWIG event generator.

4.1 Athena

Before any physics work could be done, a large amount of preparatory computing had to be completed. This involved setting up an area on lxplus, one of the CERN computing services. Then the Athena package had to be 'built', to ensure it had all the requisite tools and input files at its disposal. Once completed, the whole collection of files is compiled and linked, and then the program athena takes a user-written `AtlfastOptions.txt` file to tell it what MC to use, how many events to generate, what type of output file to produce and so on. In addition, there is an input file called `AtlfastStandardOptions.txt` in which the user can alter certain parameters relating to the reconstruction program, i.e. cuts on all the standard observables.

Athena produces 57 ntuples, representing a wide range of quantities, and can be made to save these in the form of ROOT histograms. These are stored in seven common blocks, according to the type of data contained. Histograms produced by athena are, at this stage, merely those made with default parameters rather than any particular processes, and have therefore not been included in this document.

4.2 NLO calculations

The software I have been using to calculate NLO cross-sections is described as a jet production package, and is a collection of FORTRAN programs that calculate parton-level cross-sections for protons, nuclei and photons (treating highly-virtual photons as hadronic objects). It performs pQCD calculations directly, rather than using the Monte Carlo method. There are eight primitive processes considered, four to LO and four to NLO. They are tabulated below:

NLO			LO			
0	$\to$	5g	0	$\to$	4g	
0	$\to$	3g2q	0	$\to$	2g2q	
0	$\to$	1g2q2Q	0	$\to$	2q2Q	
0	$\to$	1g4q	0	$\to$	4q	[3]

Obviously, these are unphysical events; they simply represent the numbers of each different type of parton involved in the reaction. No distinction is made between quarks and antiquarks, or between s-, t- and u-channel events. For example, '$0 \to 1g4q$' includes the processes $g + q \to q + q + \overline{q}$, $q + \overline{q} \to q + \overline{q} + g$ etc. LO processes involve a factor of α_s^2 in the matrix element, whilst those at NLO involve α_s^3. Only by considering all these processes can inclusive jet cross-sections be calculated. Furthermore, parton-level processes with identical initial and final states can interfere with each other. Once the parton-level final states have been calculated, the particles are hadronised into tracks and then clustered into jets using two standard jet-finding algorithms. No detector simulation is included at this stage.

Histograms representing cross-sections can be plotted against pseudorapidity (η), relative transverse momentum ($p_\perp$) and azimuthal separation of jets ($\Delta\phi$). Each quantity may be calculated three times; twice using a cone-finding algorithm, with parameter R=1.0 and R=0.7, and once with a $k_\perp$-clustering algorithm, with parameter D=1.0. Here R is the maximum solid angle separating any two tracks for them to considered part of the same jet;

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \quad ; \qquad \Delta R \leq R,$$

while D parameterises the separation of any two tracks in $p_\perp$-space. Here $k_\perp$ means minimum relative $p_\perp$ between tracks; that is, the momentum-vectors of the tracks are projected onto the $k_x - k_y$ plane, a vector is drawn to complete the triangle, and $k_\perp$ is the magnitude of this vector. The plots at the end of this report are of $p_\perp$, using the cone algorithm with D=1.0, in four pseudorapidity regions.

The user's choice of PDF is specified in the `testcard` file, along with other vital parameters, and then included in the main program in the process known as convolution. Essentially, the following calculation is performed:

$$\sigma(q^2) = \sum_{i=1}^{n_f} \int_{x_1} \sum_{j=1}^{n_f} \int_{x_2} \hat{\sigma}_{1,2\to X} f_1^i(x_1, q^2) f_2^j(x_2, q^2) dx_1 dx_2$$

in which σ is the hadronic cross-section, $\hat{\sigma}_{1,2\rightarrow X}$ is the partonic cross-section (for partons in hadrons 1 and 2 going to some final state X), $f^i_{1,2}(x_{1,2}, q^2)$ is the form-factor for partons 1 and 2, and n_f is the number of flavours being considered (typically four or five).

A number of alterations had to be made to the code, including the writing of a 'custom-made' histogram-fitting function. This was used to produce the plots included in this document. On each page two plots are shown, the left-hand one showing cross-sections calculated using the (old) 12-parameter ZEUS PDF set, and the right-hand one using the (new) 11-parameter set. The coloured bands represent errors on gluon PDF alone, whilst the solid line with vertical error bars represents total (valence, sea and gluon) errors. CTEQ5M is plotted as well to provide comparison with an 'errorless' PDF. (The recently-published CTEQ6 sets have errors [4], but these are characterised by large errors on gluon distribution at high x, and in any case have not yet been integrated into Frixione's and Ridolfi's NLO package.)

It can clearly be seen that, although gluon errors are reduced with respect to the two PDF sets in certain $p_\perp$ intervals, the total errors remain large. (The egregiously large total-error spikes in the first three pairs of plots are almost certainly artefacts of the histogramming algorithm, and should not be treated as either physically significant or faults in the NLO program. Further investigation will be required to remedy this.)

5 Plans for future work

One of the main long-term goals for my project, as mentioned briefly above, is to use PDFs based on ZEUS data to constrain discovery potential for a number of physics processes expected to occur at LHC, as 'seen' by the ATLAS experiment. Using PDFs with both statistical and systematic errors will be a major step forward, as it will provide a realistic benchmark against which to compare data from ATLAS when the experiment begins. Using a Monte Carlo generator with fully integrated NLO calculations will hopefully facilitate the generation of high-statistics data with parton showering (which NLO programs don't do [5].) It may also be desirable to calculate cross-sections for background events (standard-model processes which can mimic the signatures of new physics phenomena, and have to be carefully disentangled from them).

6 References

[1] Particle data group: http://www-pdg.lbl.gov/2002/strucfunrpp.ps

[2] ATLAS software page: http://atlas.web.cern.ch/Atlas/GROUPS/SOFTWARE/OO/sit/

[3] S.Frixione and G.Ridolfi in `jetdoc.txt`

[4] G.Latino, 'Jet Physics at TeVatron': hep-ex/0401/0401020

[5] J.Campbell, 'Collider Physics at NLO and the Monte Carlo MCFM':
http://www.hep.anl.gov/johnmc/talks/putalk.pdf

7 Acknowledgements

I would like to thank Chris Targett-Adams for his ongoing help with the NLO code, without which this project would scarcely have left the ground; and Mark Lancaster, Angela Wyatt and Stewart Boogart for general assistance.

8 Plots

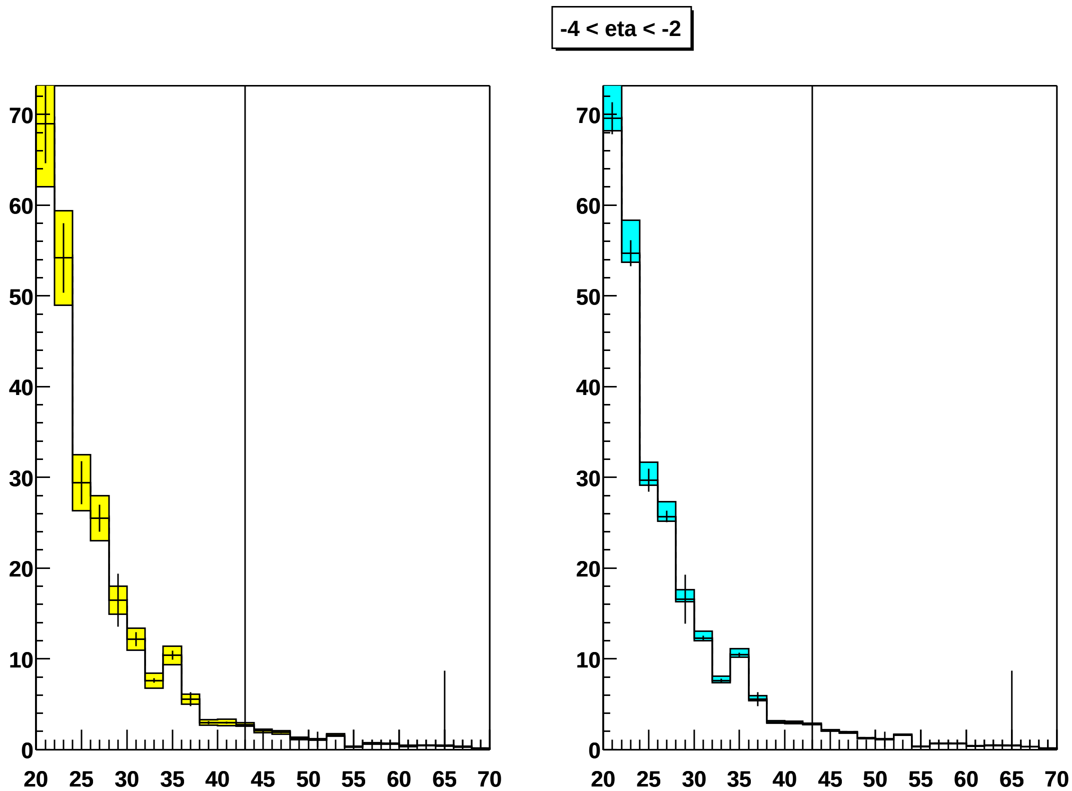

Figure 3: Transverse momentum against cross-section in the far backwards region: left-hand plot shows old (12-parameter) ZEUS fit, right-hand plot shows new (11-parameter) plot. Dashed line shows CTEQ5M for comparison, solid line with crosses shows ZEUS fit with total errors, colour shows ZEUS fit with gluon errors only. Units are μb and GeV.

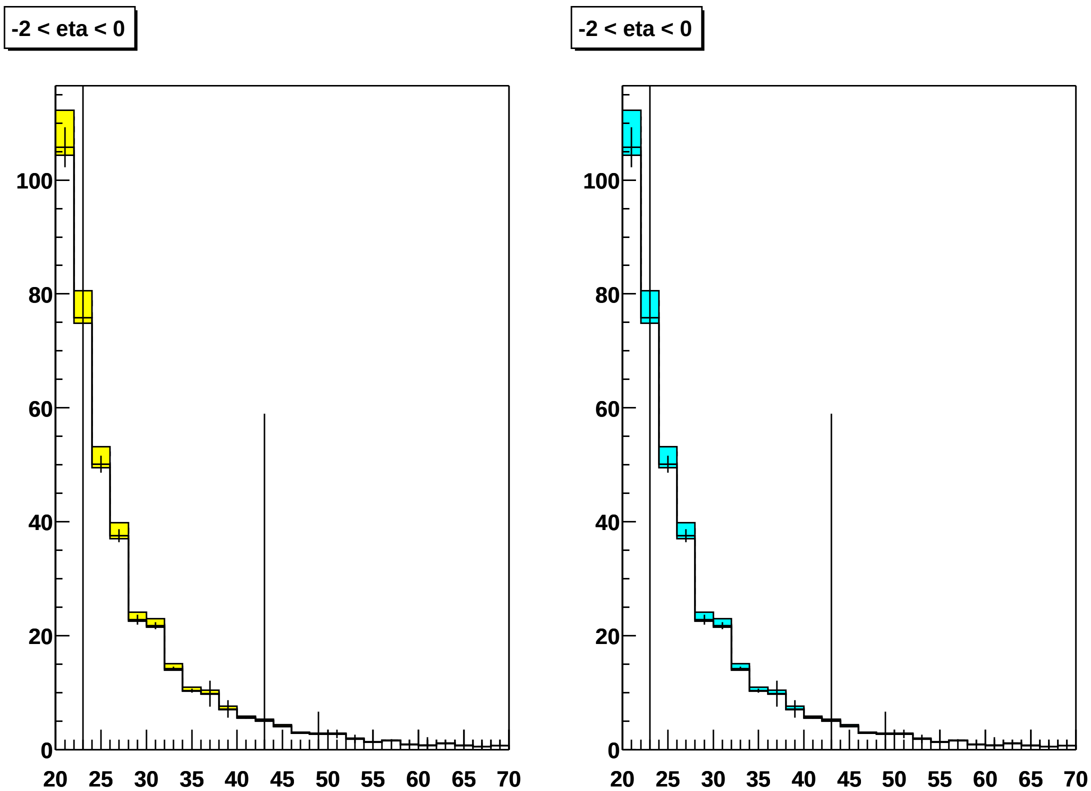

Figure 4: Transverse momentum against cross-section in the mid backwards region: left-hand plot shows old (12-parameter) ZEUS fit, right-hand plot shows new (11-parameter) plot. Dashed line shows CTEQ5M for comparison, solid line with crosses shows ZEUS fit with total errors, colour shows ZEUS fit with gluon errors only. Units are μb and GeV.

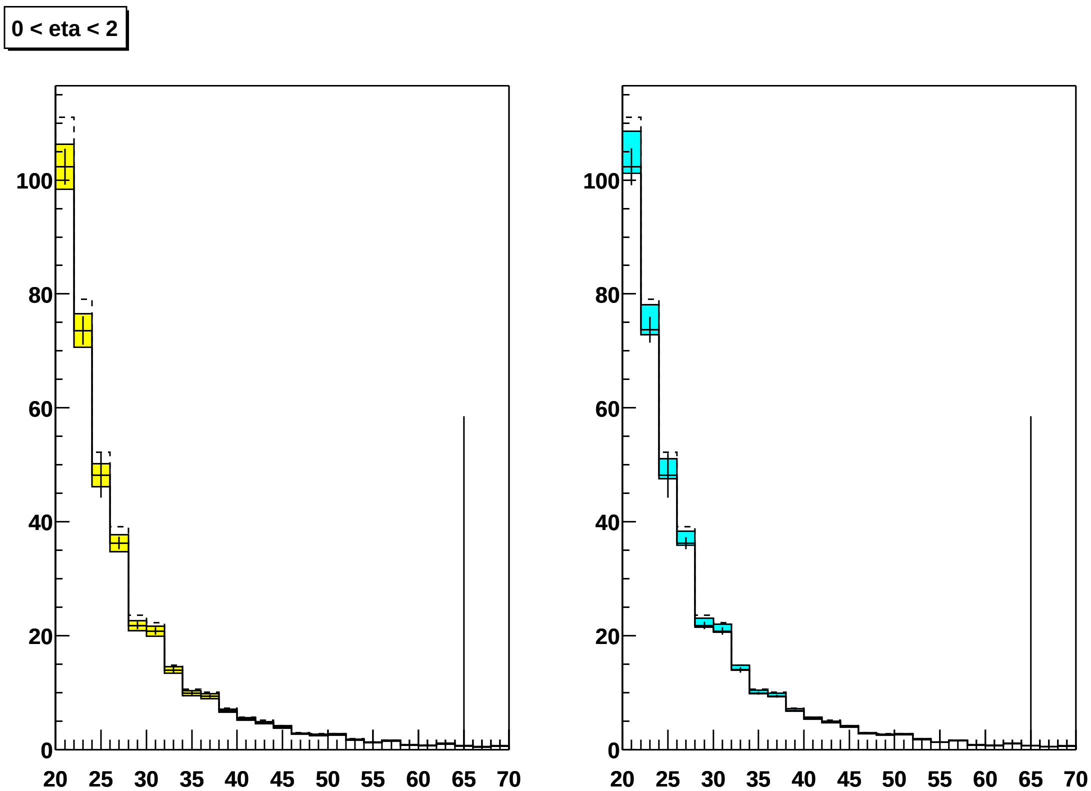

Figure 5: Transverse momentum against cross-section in the mid forwards region: left-hand plot shows old (12-parameter) ZEUS fit, right-hand plot shows new (11-parameter) plot. Dashed line shows CTEQ5M for comparison, solid line with crosses shows ZEUS fit with total errors, colour shows ZEUS fit with gluon errors only. Units are μb and GeV.

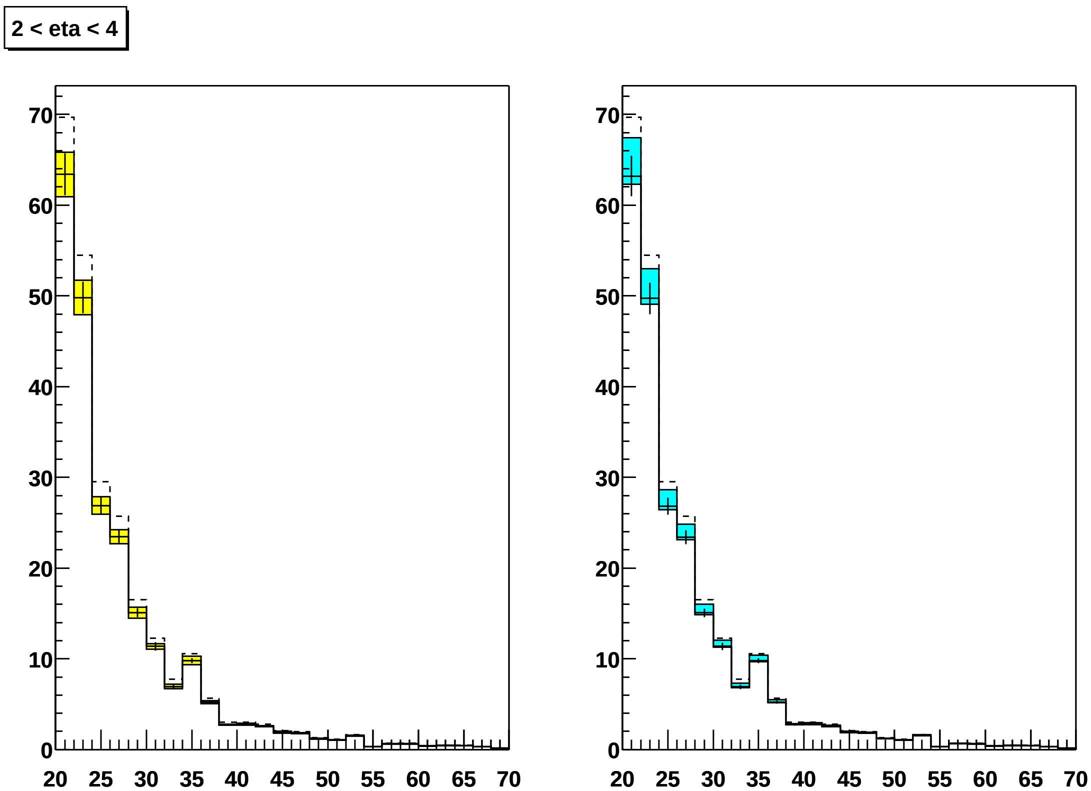

Figure 6: Transverse momentum against cross-section in the far forwards region: left-hand plot shows old (12-parameter) ZEUS fit, right-hand plot shows new (11-parameter) plot. Dashed line shows CTEQ5M for comparison, solid line with crosses shows ZEUS fit with total errors, colour shows ZEUS fit with gluon errors only. Units are μb and GeV.