D. KOSOWER

# EXTRACTING PARTON DENSITIES FROM COLLIDER DATA

Search for new physics requires a precise understanding of known physics

At high-energy hadron collider experiments,

known physics = perturbative QCD
+ non-perturbative inputs
+ small corrections

Non-perturbative inputs
$\alpha_s(M_Z^2)$
parton distribution functions
[fragmentation functions]

Inputs to theoretical predictions

or outputs of experimental measurements

Goal: 1% measurement

NLO calculations are a bare minimum

## $p\bar{p}$ collisions

| | |
|---|---|
| Pure jet production | • Search for higher-dimension operators (compositeness, heavy colored particles, ...) |
| | • Background to top in $t\bar{t} \to 6$ jets (w/ b tag) |
| Heavy quark, meson, onium production | • Measurement of parton distribut functions and fragmentation funct. |
| Photon + jet production | |
| W, Z + jet production | • Background to top in $t\bar{t} \to W + 4$ jets, $W \to \ell \not{E}_T$ |
| | • Precision measurement of W m. |
| | • Missing $E_T$ searches (Z', susy |

## $e^+e^-$ collisions

| | |
|---|---|
| Jet production | • Precision measurement of $\alpha_s$ |
| | • Ruling out light colored fermions or scalars (neutral) |

## $ep$ collisions

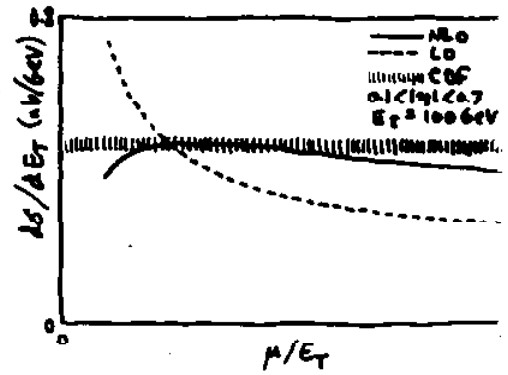
| | |
|---|---|
| Jet production | • Measurements of $\alpha_s$ & gluon distrib |

- Renormalization scale & normalization uncertainty

Physical quantities should be independent of $\mu$; but truncation to finite order in perturbation theory introduces spurious dependence

Leading order: large uncertainty from $\alpha_s(\mu)$.
Next-to-leading order: matrix element has explicit dependence on $\mu$ which compensates that in $\alpha_s(\mu)$.

→ Normalization uncertainty reduced




- Dependence on cuts & other experimental parameters

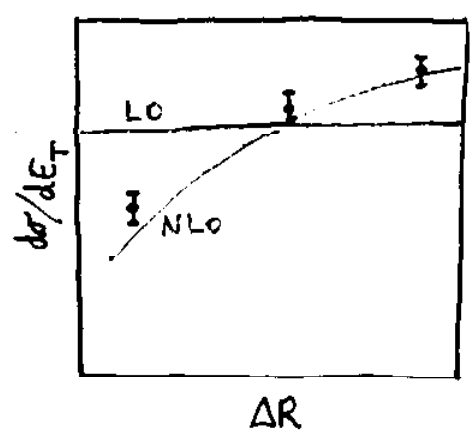
Measurements depend on jet-defining parameters such as cone size $\Delta R$ or minimum $E_T$.

Leading order: jets modelled by a lone parton, so calculated quantities lack (or have wrong) dependence on these parameters.

Next-to-leading order: real radiation inside cone & soft radiation introduce required dependence.

→ Data modelled more faithfully

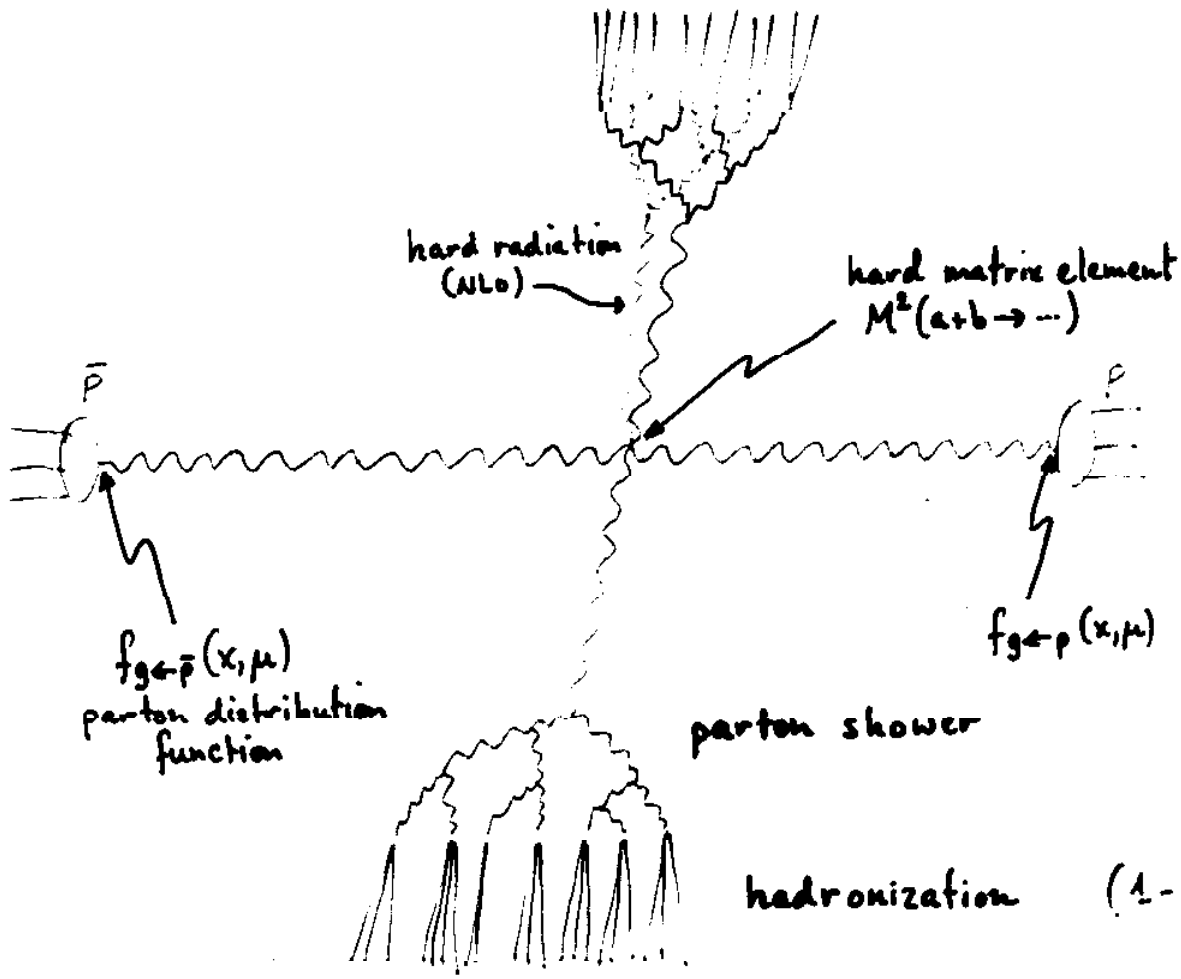
→ Applicability of perturbation understood quantitatively




In jet physics, expansion is not just in $\alpha_s(\mu)$ but in $\alpha_s(\mu)\ln^2 y_{IR}$ and $\alpha_s(\mu)\ln y_{IR}$ as well.

NLO is the first order in which these multiscale quantities are calculated quantitatively

# QCD Calculations in Jet Physics




(1 - 2 (a)

$$\sigma^{LO}(n\,\text{jets})\Big|_{cuts} = \int dx_1 dx_2 \sum_{ab} f_{a\leftarrow p}(x_1) f_{b\leftarrow \bar{p}}(x_2) \int_{cuts} dPhase(k_f)\, M^2(a+b\to\{f\})\, JetSelect(\{$$

$$\sigma^{NLO}(n\,\text{jets})\Big|_{cuts} = \int dx_1 dx_2 \sum_{ab} \Big\{ f_a(x_1) f_b(x_2)\, \hat{\sigma}^{LO}_{ab}(x_1,x_2\to n) + \alpha_s f_a(x_1) f_b(x_2) K(x_1,x_2) \otimes_p$$

(schematic)

LO

soft + collinear + virt. si

$$+\alpha_s f_a(x_1) f_b(x_2)\, \hat{\sigma}^{NLO\,finite}_{ab}(x_1,x_2\to n)$$

finite parts of virt. c
bulk of theory calculati

$$+\alpha_s f_a(x_1) f_b(x_2)\, \hat{\sigma}^{NLO\,finite}_{ab}(x_1,x_2\to n+1)$$

finite parts of real

$$+\alpha_s \big(C_a(x_1) f_b(x_2) + f_a(x_1) C_b(x_2)\big)\, \hat{\sigma}^{LO}_{ab}(x_1,x_2\to n) \Big\}$$

initial state coll. + virt. sing

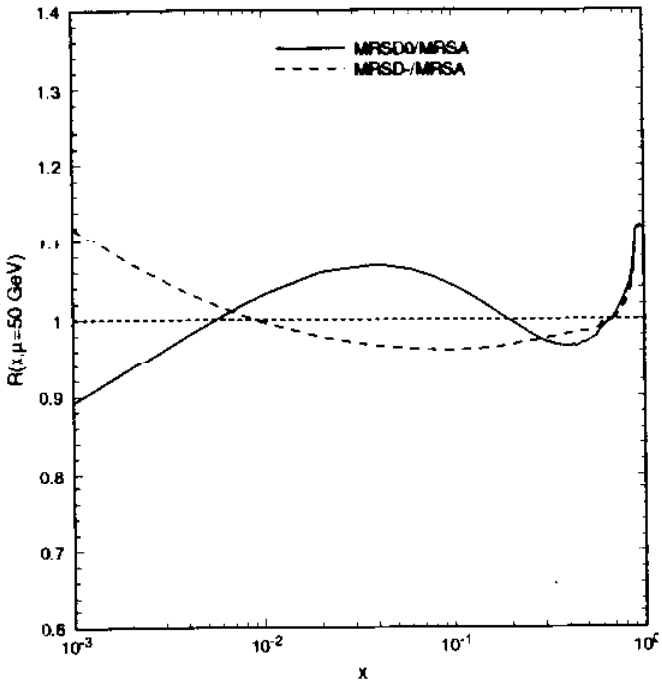



Figure 1: The ratio of the 'single effective parton density' of Eq. 4 for the $\mathrm{MRSD}_0$ and $\mathrm{MRSD}_-$ distributions compared to the MRSA parameterisation at $\mu = 50$ GeV.

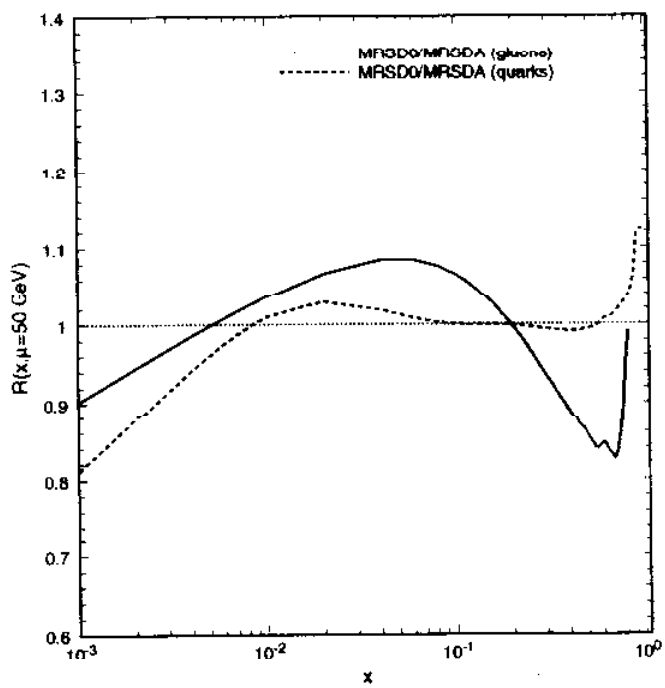
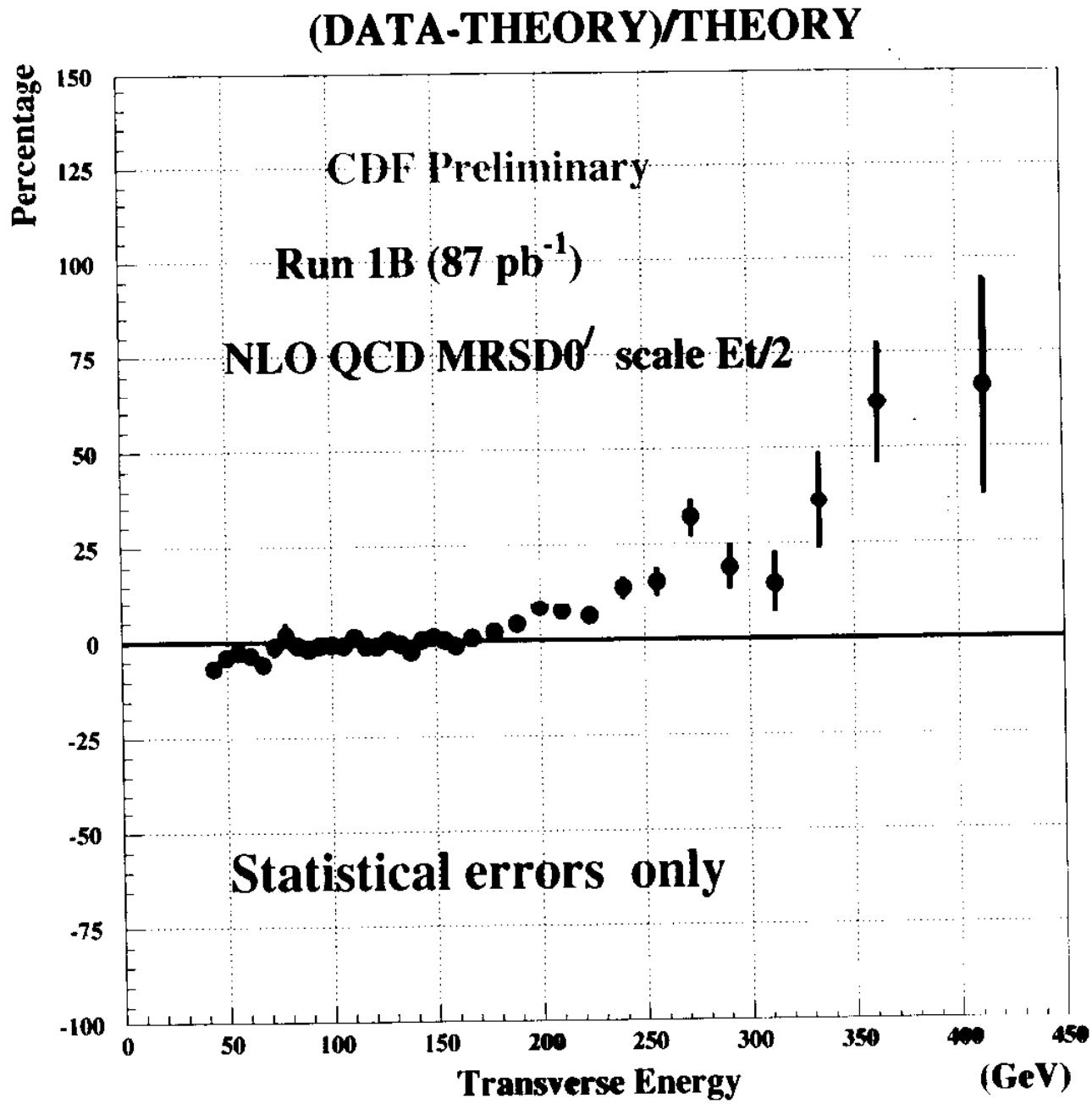
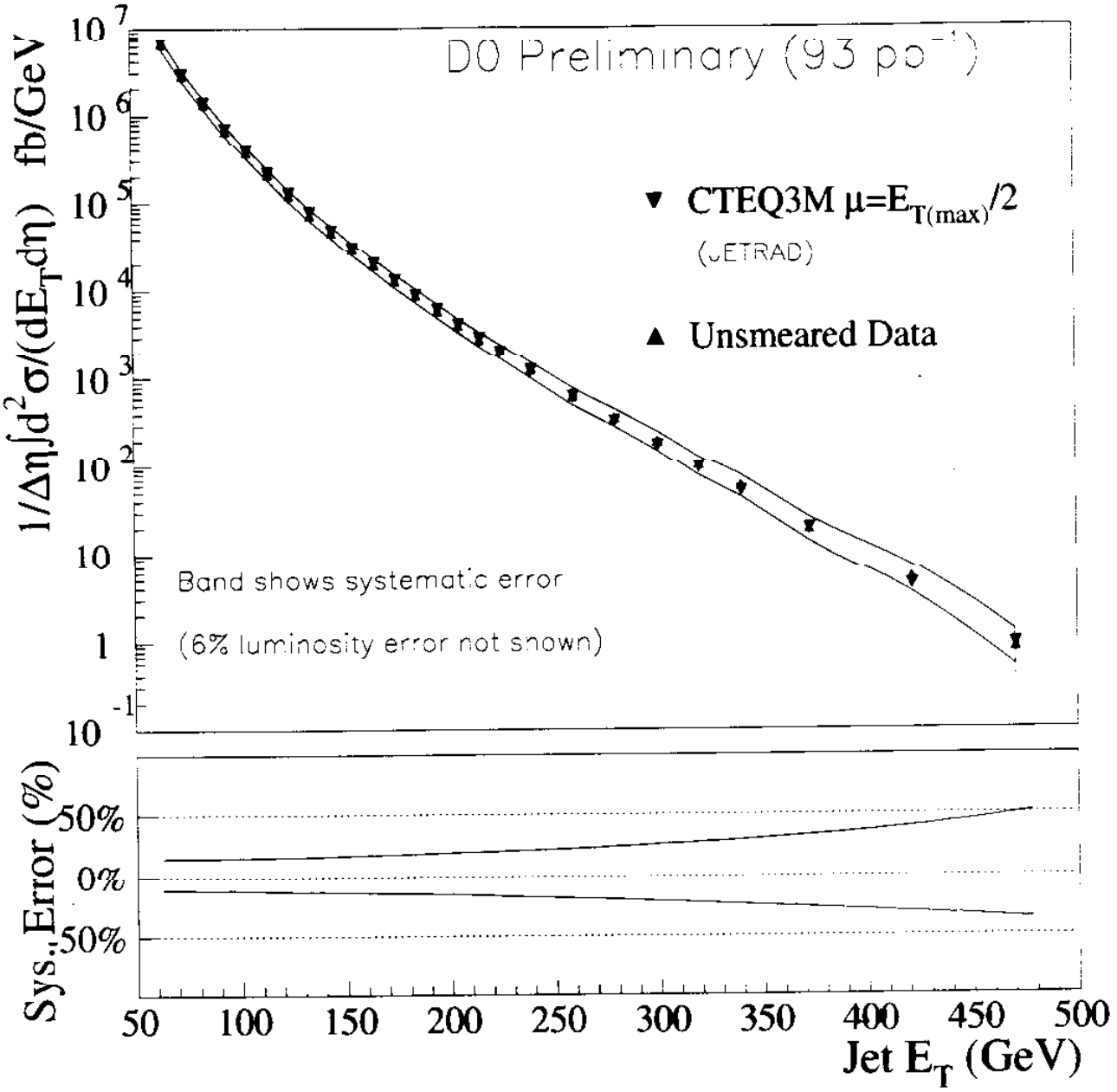



Figure 2: The ratio of the gluon and quark parton densities in the $\mathrm{MRSD}_0$ distribution compared to the MRSA parameterisation at the same scale.

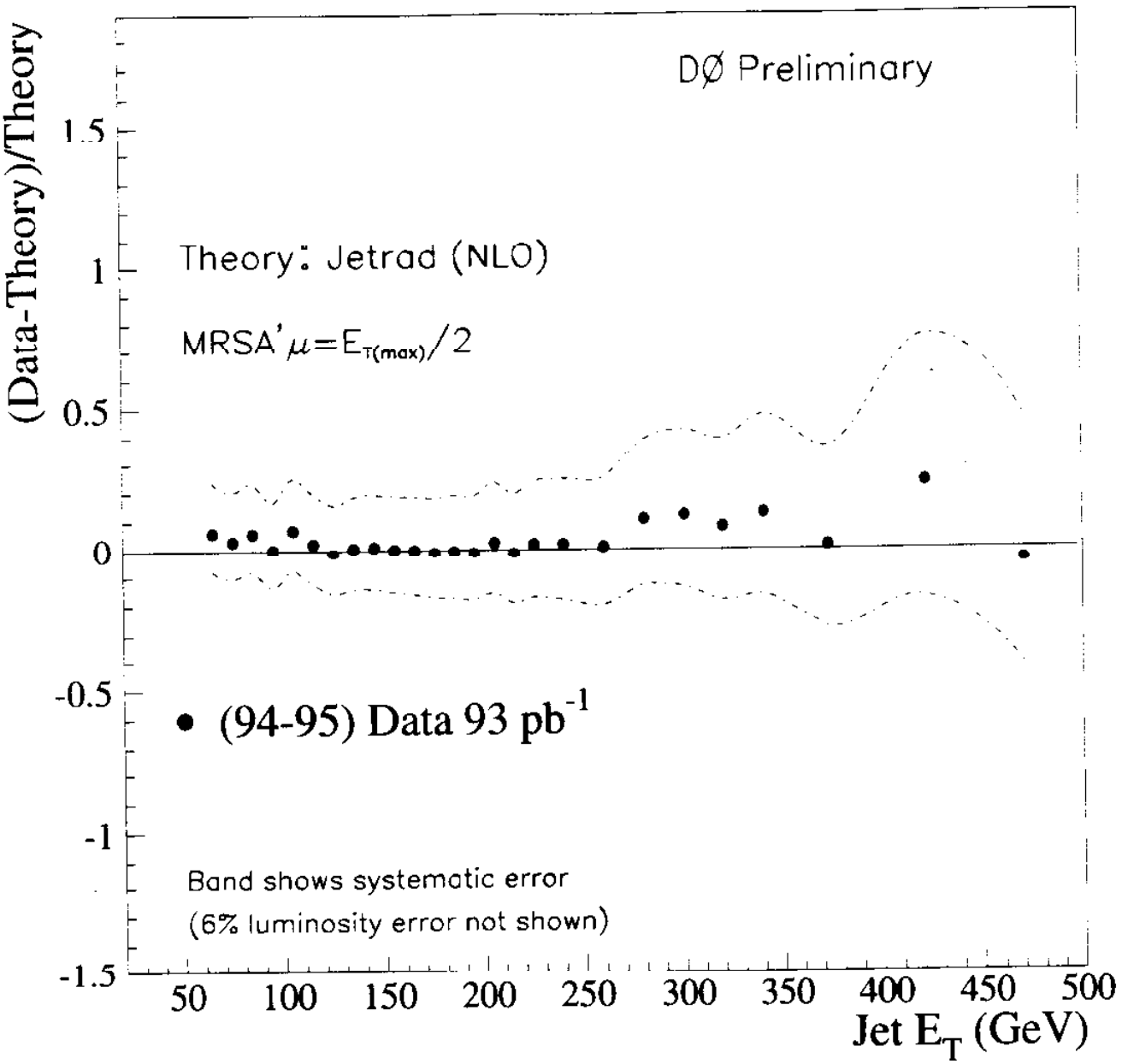



CDF http://www-cdf.fnal.gov/physics/new/qcd/
qcd=plots, inclusive_jets, ~~public~~
~~inclu~~ qcd_data_94b2000_mrsd0

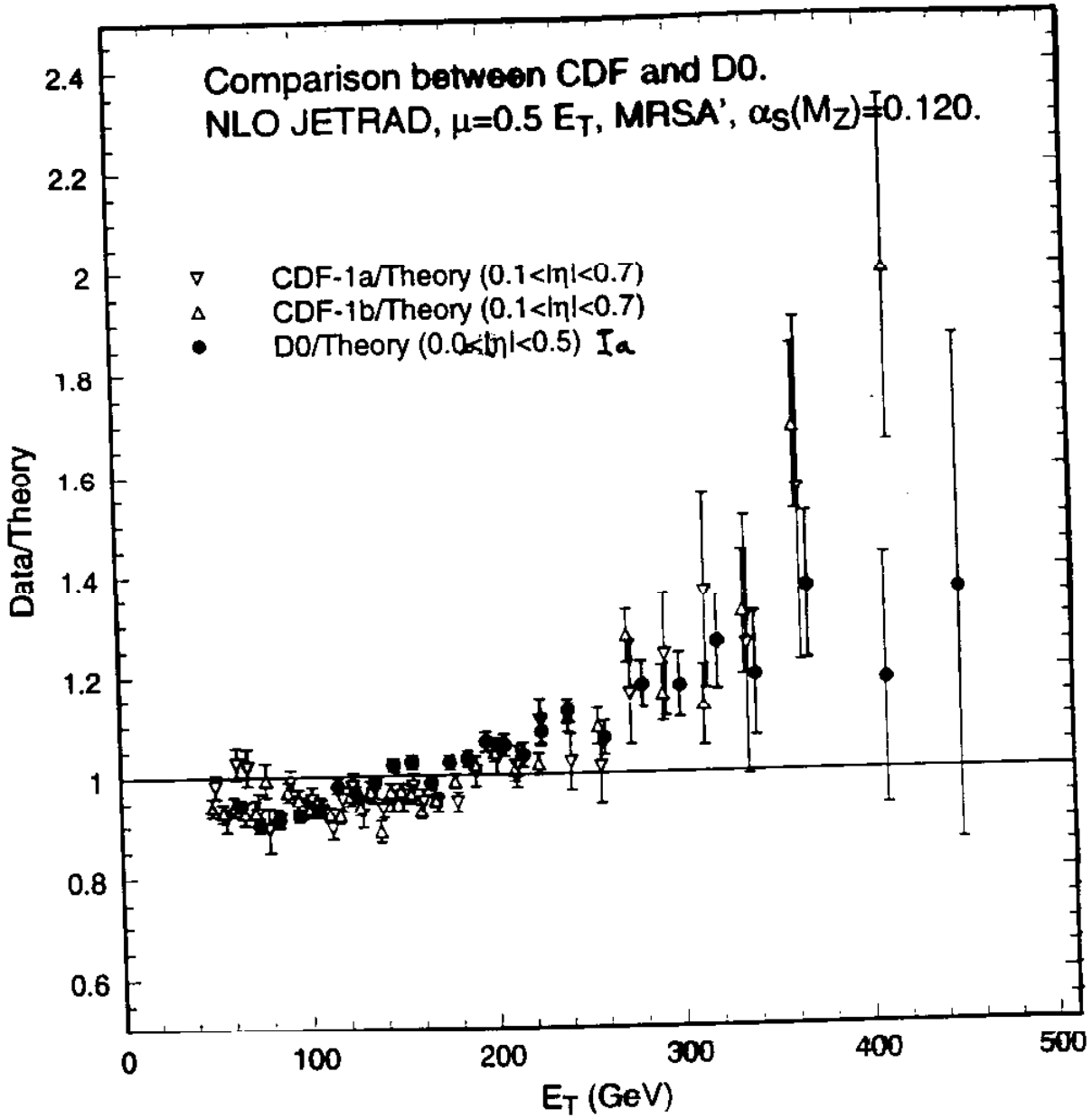
Spring '97



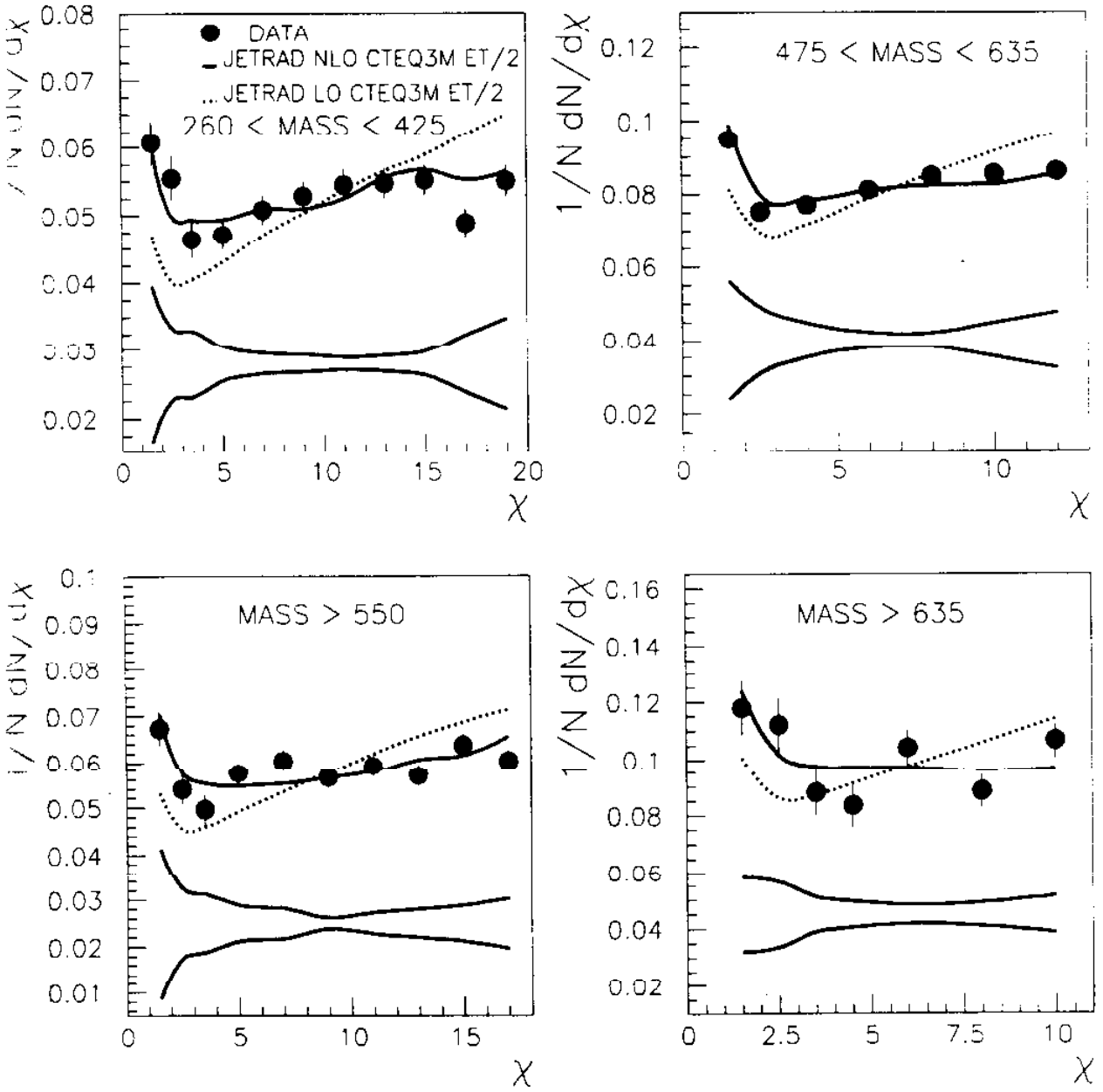

DØ, Madison, March '9




DØ, Madison, March


Comparison between CDF and D0.
NLO JETRAD, μ=0.5 E_T, MRSA', α_S(M_Z)=0.120.
CDF-1a/Theory (0.1<|η|<0.7)
CDF-1b/Theory (0.1<|η|<0.7)
D0/Theory (0.0<|η|<0.5) Ia
Data/Theory
E_T (GeV)
2.4
2.2
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0
100
200
300
400
500

# DØ PRELIMINARY




from DØ, Warsaw '96

WHAT HAVE WE LEARNED?

- WHETHER OR NOT THE CDF EXCESS TURNS OUT TO BE REAL. -

- THE GLUON DISTRIBUTION AT MODERATE & HIGH $x$ IS NOT KNOWN TO BETTER THAN 10-15
- TEVATRON DATA CAN HELP PIN IT DOWN
- TEVATRON DATA IS MORE IMPORTANT HERE THAN DIS DATA.

# HONEST ERROR ESTIMATES

- Spread of predictions using different parton distribution sets (MRS, MRS′, GRV, CTEQ, ...) is NOT an error estimate: different sets make common assumptions, that is, they are "correlated" (though in ways hard to quantify); different sets from same authors reflect different assumptions.

  Sets corresponding to $1\sigma$ deviations in $\chi^2$ (in a representative collection of directions in parameter space) are what we need to make honest and sensible error estimates. Without them, we cannot make quantitative statements about the error contribution from pdfs, or talk sensibly about discrepancies between experiments

- Variation of a prediction w/renormalization scale indicates that there is theoretical uncertainty, but it is NOT a quantitative estimate of the error. Because of logs, comparing LO and NLO generally isn't either.

  An honest error estimate requires yet higher-order calculations - NNLO.

  Leading wedge: higher-loop calculations of two-point functions,

# INCLUSIVE TWO-JET STUDIES

$$p\bar{p} \to 2 \text{ jets} + X$$

$$\sigma = \int dx_1 dx_2 \; f(x_1) f(x_2) \; \hat{\sigma}(x_1, x_2)$$

known pdf $\oplus$ theory $\longrightarrow$ prediction for experiment

data $\oplus$ theory $\longrightarrow$ measurement of pdf

Most general distribution $\dfrac{d^3\sigma}{dE_T \, d\eta_1 \, d\eta_2}$

$\updownarrow$ LO, fixed $E_T$

$\dfrac{d^2\sigma}{dx_1 dx_2}$

At LO,

$$x_{1,2} = \frac{E_T}{\sqrt{s}} \sum_{p=1,2} e^{\pm \eta_p}$$

At higher order, this correspondence is broken, but with a sensible jet algorithm, the corrections are not too large, and allow the extraction of pdf's.

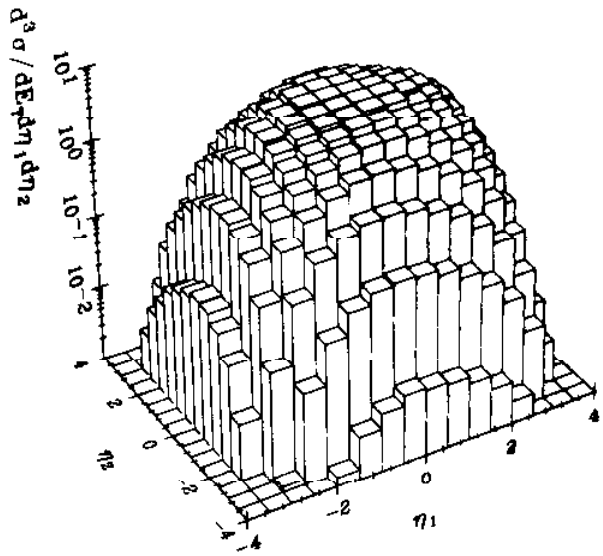



Figure 1: The next to leading order triple differential distribution for 45 GeV $< E_T <$ 55 GeV and $\mu = E_{T1}$ for $\mathrm{MRSD}_0$ parton densities.

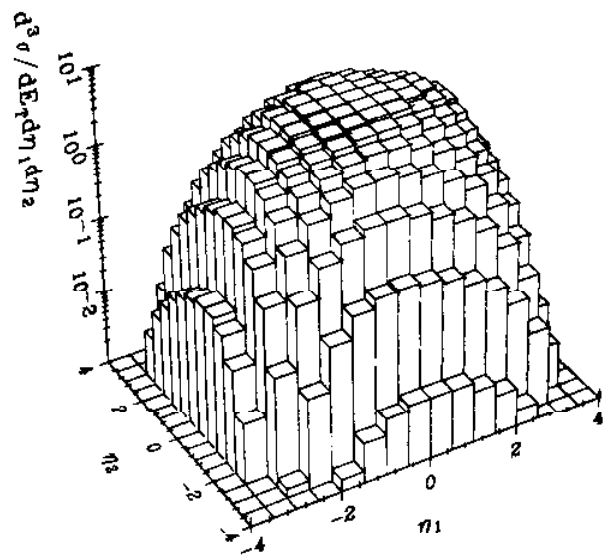
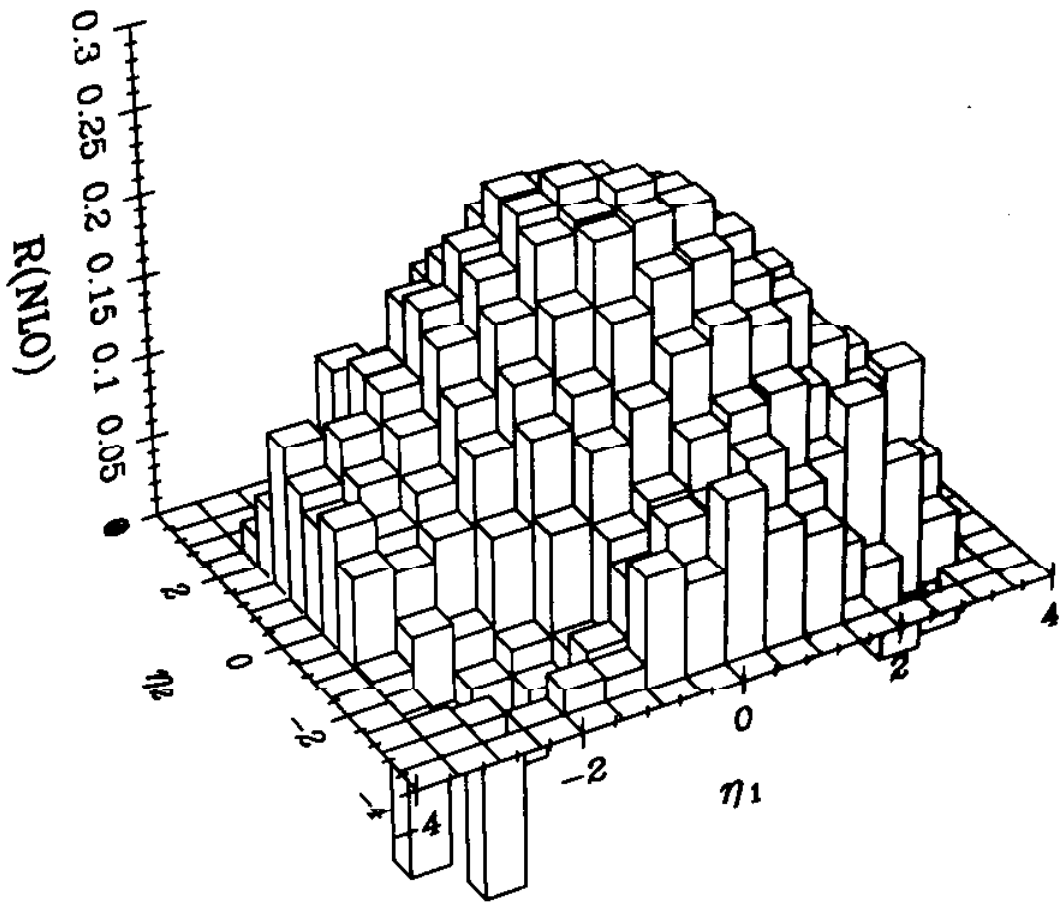
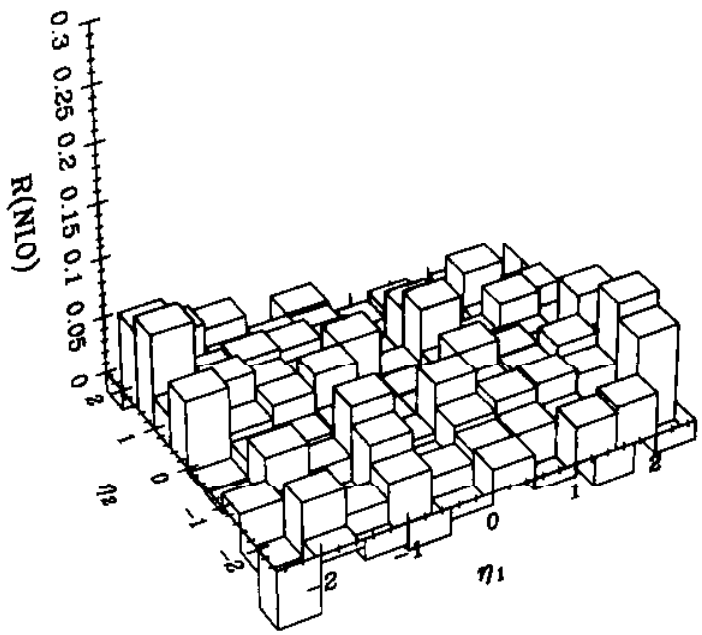



Figure 2: The next-to-leading order triple differential distribution for 45 GeV $< E_T <$ 55 GeV and $\mu = E_{T1}$ for $\mathrm{MRSD}_-$ parton densities.

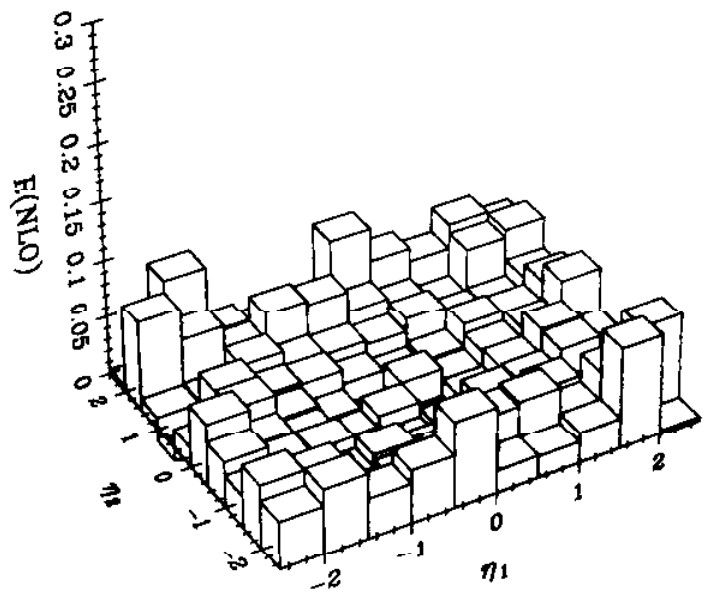

R(NLO)
0.3 0.25 0.2 0.15 0.1 0.05
$\eta_2$
2
0
-2
-4
$\eta_1$
-4
-2
0
2
4




The next to leading order prediction for the fractional ratio of the triply differential cross section at different renormalization scales,

$$\frac{(\langle d\sigma(\mu = E_{T1})\rangle - \langle d\sigma(\mu = \lambda E_{T1})\rangle)}{\langle d\sigma(\mu = E_{T1})\rangle}$$

with $\lambda = 1/2$, for 45 GeV $< E_T <$ 55 GeV and MRSD$_-$ structure functions.




The same with $\lambda = 2$.

To extract information about the parton distributions,

avoiding both theoretical & experimental problems with normalisatio

minimize a quantity like

$$\sum_{\substack{E_T,\,\eta_1\times\eta_2 \\ \text{bins}}} \left( \int_{\text{bin}} \frac{d^3\sigma^{EXP}}{dE_T\,d\eta_1\,d\eta_2} - c\int_{\text{bin}} \frac{d^3\sigma^{TH}}{dE_T\,d\eta_1\,d\eta_2} \right)^2$$

Ultimately: global fit along with all DIS data

Initially: fit only parameters poorly determined by DIS data
(gluon dist.)

# HOW SHOULD WE FIT TO DATA?

Conceptually, use the DIS model

Iterate

Compute distribution (e.g. $F_2$)
↓
Compare with data
↓
Adjust parameters

known analytically using moments directly in terms of parameters in pdfs

with $f_p(x, Q_0^2) = A_p \, x^{-\lambda_p} (1-x)^{\beta_p} (1 + \epsilon_p \sqrt{x} + \delta_p x)$

but for jet distributions this is s l o o o o w...

because it involves re-running a jet program at every iteration.

Need a better way — and one should be possible, because most of the computation — cuts, clustering, jet algorithm — doesn't really involve the pdfs in a deep way.

Another approach that's too slow: generating values at points on a grid of parameter values. 15 parameters → $n^{15}$ points for $n$ values/parameter.

# A LIMITED FORMALISM FOR DIS

Graudenz, Hampel, Vogt, & Berger [hep-ph/9506333]

Mellin transform for non-factorizing integrals

Bin in $(x_B, Q^2) \rightarrow \frac{d\sigma}{dx_B\, dQ^2}$

Evaluate $\int_a^1 dx_B \frac{d\sigma}{dx_B\, dQ^2}$ using (complex) Mellin moments

- Does not allow $\mu^2_{R,F}$ to depend on phase-space variables — essential in $p\bar{p}$
- Their intermediate forms discouraged them from using an efficient contour

# GLUEBALL GLUEBALL SCATTERING
## in quarkless QCD

The n-jet cross section is

$$\sigma^{LO}(n\text{jets})\Big|_{\text{cuts}} = \int dx_1\, dx_2 \int d\text{Phase}\,(x_1 k_G + x_2 k_{G'} \to \{k_i\}_1^n)$$

$$\times\; f_{g\leftarrow G}(x_1, \mu_F^2(\{k_i\}, x_{1,2}))\; f_{g\leftarrow G}(x_2, \mu_F^2(\{k_i\}, x_{1,2}))$$

$$\times\; \alpha_s^n(\mu_R^2(\{k_i\}, x_{1,2}))\; \hat{\sigma}(gg \to \{k_i\})\; \text{JetSelect}(\{k_i\})$$

The renormalization & factorization scales depend on the final-state momenta (and thus on $x_{1,2}$).

Write the gluon distribution in terms of its Mellin transform,

$$f_{g\leftarrow G}(x, \mu^2) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} dz\; x^{-z}\; f^z_{g\leftarrow G}(\mu^2)$$

The contour runs to the right of all singularities of $f^z_{g\leftarrow G}$

$f(x,\mu^2)$ satisfies the Altarelli-Parisi equation,

$$\frac{d f_{g\leftarrow G}(x,\mu^2)}{d\mu^2} = P(x,\mu^2) \bullet f_{g\leftarrow G}(x,\mu^2)$$

↳ convolution

which has a simple solution for its Mellin transform

$$f^{z}_{g\leftarrow G}(\mu^2) = E^{z}(\alpha_s(\mu^2), \alpha_s(\mu_0^2))\, f^{z}_{g\leftarrow G}(\mu_0^2)$$

↳ evolution operator

Therefore

$$f_{g\leftarrow G}(x,\mu^2) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} dz\, x^{-z}\, E^{z}(\alpha_s(\mu^2), \alpha_s(\mu_0^2))\, f^{z}_{g\leftarrow G}(\mu_0^2)$$

- contains all fit parameters except $\alpha_s(\mu_0^2)$
- independent of all integration variables except $z$.

So each step of a fit procedure requires computing only the contour integrals,

$$\sigma^{LO}(n\text{ jets})\Big|_{cuts} = -\frac{1}{4\pi^2}\int_{c-i\infty}^{c+i\infty} dz_1 \int_{c-i\infty}^{c+i\infty} dz_2\, f^{z_1}_{g\leftarrow G}(\mu_0^2)\, f^{z_2}_{g\leftarrow G}(\mu_0^2)\, \Sigma^{z_1}$$

precomputed

$\alpha_s(\mu_0^2)$, last fit parameter

$$\Sigma^{z_1,z_2} = \int dx_1\, dx_2 \int d\mathrm{Phase}\;\; x_1^{-z_1}\, E^{z_1}(\alpha_s(\mu^2), \alpha_0)\, x_2^{-z_2}\, E^{z_2}(\alpha_s(\mu^2), \alpha_0)$$

$$\times\; \alpha_s(\mu^2)\, \hat{\sigma}(gg \to \{k_i\})\; \mathrm{JetSelect}(\{k_i\})$$

For $\alpha_0$: generate a set of $\Sigma^{z_1,z_2}$ for different $\alpha_0$, & interpolate.

# CONTOUR INTEGRALS

How should we do the contour integrals?

Study a similar problem, the evolution of parton distributions,

$$f(x,\mu^2) = \underbrace{\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} dz}_{\text{compute numerically}} \; \underbrace{x^{-z}\, E^{z}(\alpha_s(\mu^2),\alpha_0)}_{\text{known analytically}} \; \underbrace{f^{z}(\mu_0^2)}_{\text{specified analytically}}$$

Textbook contour: parallel to the imaginary axis.

$x^{-z}$ purely oscillating

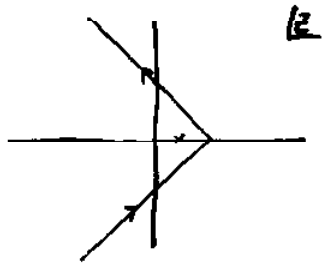
$f^{z}(\mu_0^2)$ has power decay: $x^{\alpha}(1-x)^{\beta}$ gives

$f^{z} \sim z^{-\beta-1}$ as $z \to \infty$

Terrible choice

Want a contour for which $z$ has an increasingly negative real part $\to$ exponential convergence from $x^{-z}$

For example, GRV use




Better choice

But: why this one?

Best choice: presumably contour of steepest descent

- Evolution is slow, so use $E^z = 1$ to find the contour
- A parabolic contour (as it turns out) gives an excellent approximation to the best contour
- Function has very nearly an exponential behavior along this contour

Procedure

Take $F(z) = x^{-z} f^{z}(\mu_0^2)$.

Find minimum of $F$ along the real axis $(c_0)$.

Compute $c_2 = \sqrt{\dfrac{2F(c_0)}{F''(c_0)}}$, $c_3 = \dfrac{F^{(3)}(c_0)}{3F''(c_0)}$

The desired contour is $z(u) = c_0 + \left(i\sqrt{u} + c_2 c_3 \frac{u}{2}\right) c_2$

Use a generalized Gauss-Laguerre quadrature to evaluate the integral

$$f(x,\mu^2) = \frac{c_2}{2\pi}\int_0^\infty \frac{du}{\sqrt{u}} e^{-u} \operatorname{Re}\Big[\underbrace{e^{u}\, x^{-z(u)} E^{z(u)}(\alpha_s(\mu^2),\alpha_0) f^{z(u)}(\mu_0^2)}_{\text{well-approximated by a low-order polynomial}}\Big]$$

$$= \frac{c_2}{2\pi}\sum_{j=1}^{n} w_j \operatorname{Re}\Big[e^{u_j}\, x^{-z(u_j)} E^{z(u_j)}(\alpha_s(\mu^2),\alpha_0) f^{z(u_j)}(\mu_0^2)\Big]$$

$u_j$ – zeros of $L_n^{(-1/2)}(u)$, $w_j$ GL weights

# BEYOND TOYS

- Add quarks: sprinkle indices
- LO → NLO: treat each contribution analogously
- Contour in DIS: use $F = f^{z}_{a \leftarrow p}(\mu_0^2) \int dE_T \, S(E_T) \frac{d\Sigma^{z}_{a}}{dE_T}(\mathrm{LO})$

  (where $S(E_T)$ compensates for the rapid fall-off in $E_T$) as a contour-determining function
- Contours in $p\bar{p}$: use stationary phase condition to obtain approximations to joint contours $z_1(t_1, t_2)$, $z_2(t_1, t_2)$.
  $[\mathrm{Im}\, F(z_1, z_2) = 0.]$

Define

$$F(z_1, z_2) = f^{z_1}_{a\leftarrow p}(Q_0^2) f^{z_2}_{b\leftarrow \bar{p}}(Q_0^2) \int dX\ S(X) \frac{d\Sigma^{z_1,z_2}_{ab}}{dX} \ ,$$

The observable we want to calculate is

$$-\frac{1}{2\pi^2} \int_0^{c+i\infty} \int_0^{c+i\infty} \mathrm{Re}\left[dz_1\, dz_2\ F(z_1, z_2) - dz_1\, d\bar{z}_2\ F(z_1, \bar{z}_2)\right]$$

Find the minimum of $F(z_1, z_2)$ for real $z_{1,2}$, label its coordinates $(c_1, c_2)$.

Parametrize the surface using two variables $t_{1,2}$,

$$z_1 = c_1 + it_1 + a_{11}t_1^2 + a_{12}t_1 t_2 \ ,$$
$$z_2 = c_2 + it_2 + a_{22}t_2^2 + a_{21}t_1 t_2 \ ,$$

Expand the equation $\mathrm{Im}\, F(z_1(t_1, t_2), z_2(t_1, t_2)) = 0$ to obtain

$$a_{11} = \frac{F^{(3,0)}}{6F^{(2,0)}}$$

$$a_{12} = -\frac{F^{(0,3)}}{6F^{(0,2)}} + \frac{1}{6J}\left[F^{(0,3)}F^{(2,0)} - 3F^{(1,2)}F^{(1,1)} + 3F^{(2,1)}F^{(0,2)} - \frac{F^{(3,0)}F^{(0,2)}F^{(1,1)}}{F^{(2,0)}}\right]$$

$$a_{21} = -\frac{F^{(3,0)}}{6F^{(2,0)}} + \frac{1}{6J}\left[F^{(3,0)}F^{(0,2)} - 3F^{(2,1)}F^{(1,1)} + 3F^{(1,2)}F^{(2,0)} - \frac{F^{(0,3)}F^{(2,0)}F^{(1,1)}}{F^{(0,2)}}\right]$$

$$a_{22} = \frac{F^{(0,3)}}{6F^{(0,2)}}$$

where

$$F^{(j_1,j_2)} = \left[\frac{\partial^{j_1+j_2} F(z_1, z_2)}{\partial z_1^{j_1} \partial z_2^{j_2}}\right]\Bigg|_{(z_1,z_2)=(c_1,c_2)}$$

$$J = \det\begin{pmatrix} F^{(2,0)} & F^{(1,1)} \\ F^{(1,1)} & F^{(0,2)} \end{pmatrix}$$

We expect the function to go like $F(c_1, c_2)e^{-g(t_1,t_2)}$, with

$$g(t_1, t_2) = \frac{F^{(2,0)}}{2F(c_1, c_2)} t_1^2 + \frac{F^{(1,1)}}{F(c_1, c_2)} t_1 t_2 + \frac{F^{(2,0)}}{2F(c_1, c_2)} t_1^2$$

which suggests the change of variables

$$t_1 = \sqrt{\frac{F(c_1, c_2)(\Delta + \Lambda)}{\lambda_+ \Lambda}}\sqrt{u_1} + \sqrt{\frac{F(c_1, c_2)(\Delta - \Lambda)}{\lambda_- \Delta}}\sqrt{u_2}$$

$$t_2 = -\sqrt{\frac{F(c_1, c_2)(\Delta - \Lambda)}{\lambda_+ \Delta}}\sqrt{u_1} + \sqrt{\frac{F(c_1, c_2)(\Delta + \Lambda)}{\lambda_- \Delta}}\sqrt{u_2}$$

where

$$\Delta = F^{(2,0)} - F^{(0,2)},$$

$$\lambda_+ = \frac{1}{2}\left[F^{(2,0)} + F^{(0,2)} \pm \sqrt{\left(F^{(2,0)} + F^{(0,2)}\right)^2 + 4\left(F^{(1,1)}\right)^2}\right]$$

are the eigenvalues of the Hessian matrix at $(c_1, c_2)$, and

$$\Lambda = \lambda_+ - \lambda_- .$$

# SUMMARY

- Want better measurements of the gluon distribution inside the proton.
- Collider data can play an important role.
- NLO theory for cross-sections exists.
- Formalism for efficient fitting.