Uncertainties of Parton Distributions and Their Implication on Physical Predictions

D. Stump, J. Pumplin, R. Brock, D. Casey, J. Huston, J. Kalk, W.K. Tung

Department of Physics and Astronomy
Michigan State University, USA

(Abstract 91, hep-ph/0006148)

New Global Analysis of DIS Data and the Strange Sea Distribution

V. Barone

Università "A. Avogadro" & INFN, Sezione di Torino, Italy

C. Pascaud and F. Zomer

LAL, IN2P3-CNRS and Université de Paris-Sud, Orsay, France

(abstract 322: hep-ph/0004268 &
Eur.Phys.J. C12, 243, 2000)

Results of the MSU Group on Uncertainties of PDFs and Phys. Observables

(based on apparatus of CTEQ global analysis)

Lagrange Multiplier Method
– for specific physical variables, say X –

- Uncertainty of X: ΔX – within "one s.d." allowed by global fit to all pertinent data sets;
- Three sets of PDFs, with $\Delta X = -1, 0, 1$, which can be used for relevant applications;
- This can be done for any X: e.g.
  · $\sigma_{W,Z}$, moments of y and p_T distributions;
  · various Parton Luminosities at a given $\sqrt{s}$;
  · σ_{Higgs}, $\cdots$

Hessian Method for Parton Distributions and for general physical variables

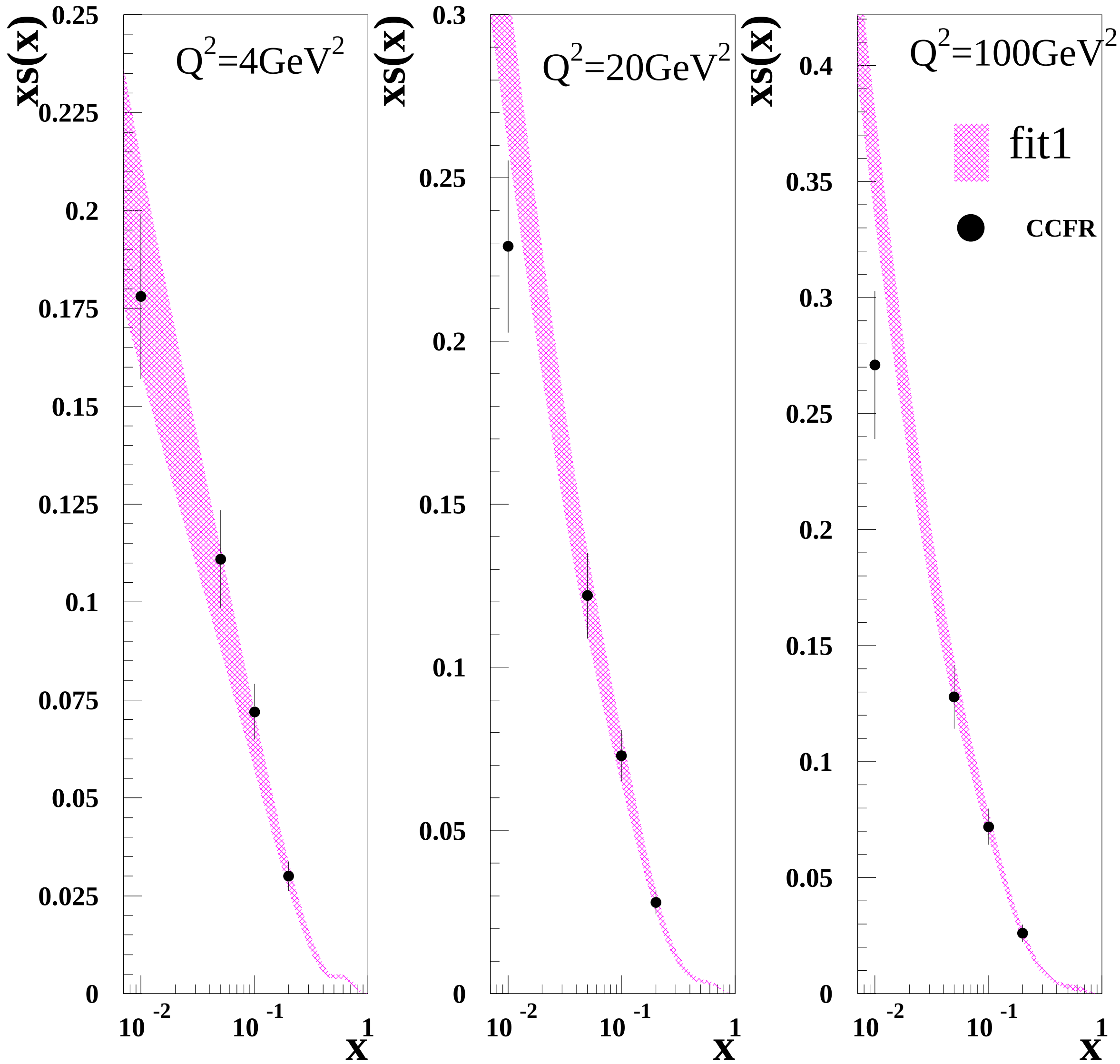
- Eigenvalues and Eigenvectors of the Hessian;
- Given physical variable X and the vector $\partial X/\partial a_i$, can calculate $\Delta X = \sqrt{2\sum_{i,j}\frac{\partial X}{\partial a_i}(H^{-1})_{ij}\frac{\partial X}{\partial a_j}}$;
- Systematically identify "Flat directions" where parametrization can be tightened, or improvements in experimental constraints are needed;
- Collection of PDFs on the surface of "one s.d." of the Hessian, suitable for probing uncertainties of a variety of physical variables.

Results from
Barone, Pascaud, and Zomer

A New Global Analysis Of Dis Data
And The Strange Sea Distribution

- A new global analysis of DIS data, characterized by an enlarged neutrino and antineutrino data set (BEBC, CDHS, CDHSW), but no CCFR;
- Special emphasis is given to the strange sector; The strange sea distribution is determined independently of the non-strange sea;
- The possibility of a charge asymmetry, $s(x) \neq \bar{s}(x)$, is tested.

Strange Sea Determination

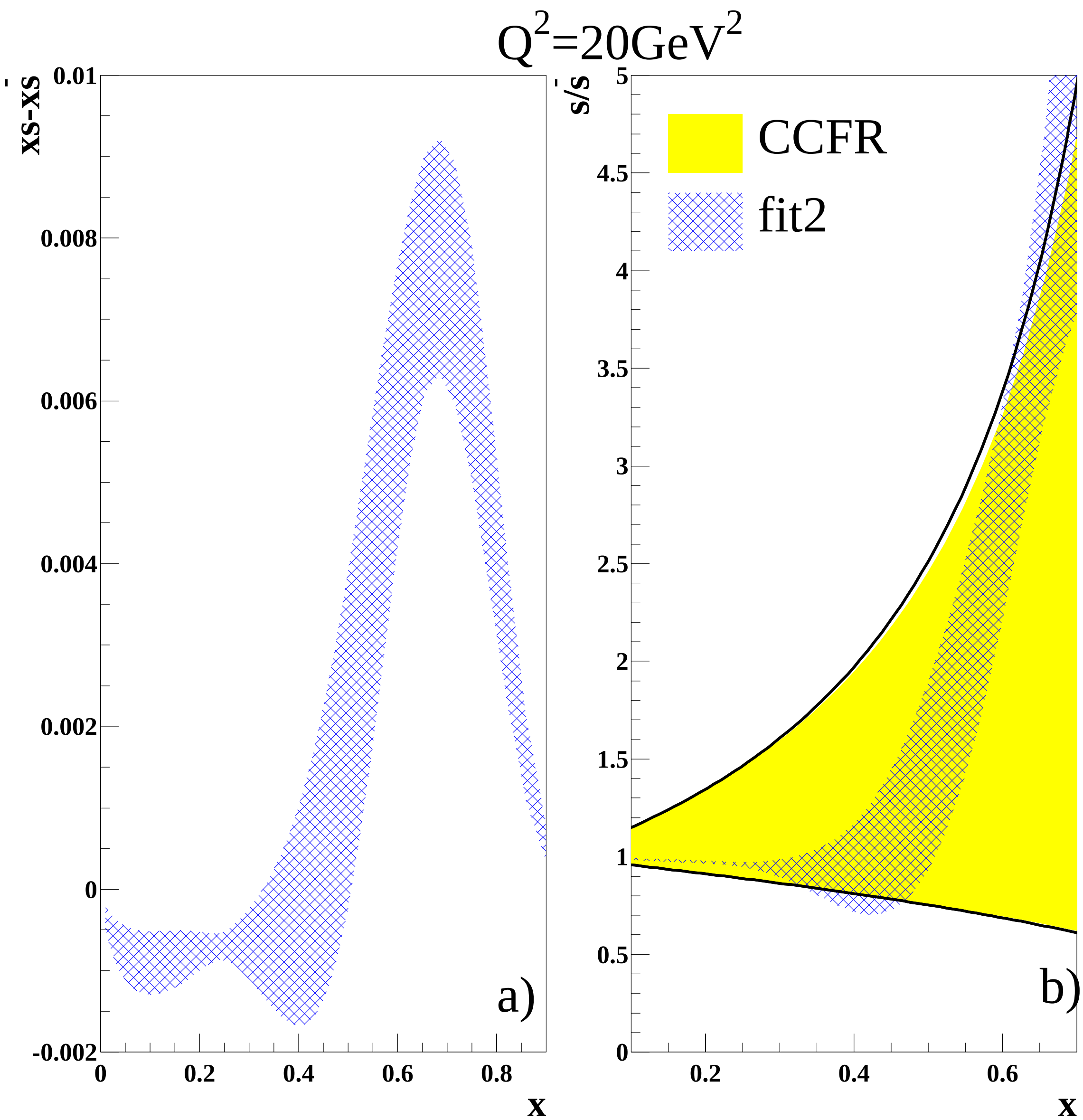


Band: uncertainty of one s.d. in combined fit, defined as $\Delta\chi^2 = 1$.

Points: CCFR dimuon determination of $xs(x, Q)$.

Barone etal.

Charge Asymmetry of Strange Sea
$x(s(x,Q) - \bar{s}(x,Q))$ and $(s(x,Q)/\bar{s}(x,Q))$



* $s(x)$ appears to be harder than $\bar{s}(x)$ – favoring the idea of intrinsic sea theory.

Barone etal.

Global QCD Analysis of Parton Distributions

In principle

- ⇒ Experimental data on all available hard scattering processes
- ⇒ NLO QCD Hard-cross-sections (or beyond) for these processes
- ⇒ Parametrized functions for the non-perturbative initial parton distributions
  + NLO QCD-evolution of these functions

Global analysis: compare theoretical calculation (based on *factorization theorems* of PQCD) to world's experimental data, to determine the fundamental QCD constants ($\Lambda_{QCD}, m_i; i =$ quark flavors) and the non-perturbative PDF parameters (usually by χ^2 minimization).

In Practice

Many subtleties and complications ...
due to imperfect theory and/or experiment,
which are under continuous development.

⇒ Several ongoing global analysis efforts. ⇐

Global QCD Fit

* Parametrization of the non-perturbative PDFs: (at $Q_0 = 1$ GeV)

$$f_i(x, Q_0) = a_0^i x^{a_1^i} (1-x)^{a_2^i} (1 + a_3^i x^{a_4^i}).$$

(with exception of $\bar{d}/\bar{u}$)

* The fitting is done by minimizing a global "chi-square" function, χ^2_{global}. This function serves as a *figure of merit* of the quality of the global fit,

$$\begin{aligned} \chi^2_{\text{global}} &= \sum_n \sum_i w_n \left[\left(N_n d_{ni} - t_{ni} \right) / \sigma^d_{ni} \right]^2 \\ &+ \sum_n \left[\left(1 - N_n \right) / \sigma^N_n \right]^2 \end{aligned} \quad (1)$$

d_{ni} : data point
σ^d_{ni}: combined error
t_{ni}: theory value (dependent on $\{a_i\}$)
for the i^{th} data point in the n^{th} experiment. w_n: a priori weighing factor for certain expts.

* This approximate χ^2 function does not have the full probabilistic significance of rigorous statistical analysis. It can be improved (cf. below).
The methods of uncertainty analysis remains valid.

Physical processes and experiments *

DIS – Neutral Current (e,μ on p,d)

SLAC, BCDMS, NMC, E665, H1, ZEUS

DIS – Charged Current ($\nu, \bar{\nu}$ on nucleus)

CCFR(F_2, F_3)

Drell-Yan – continuum (lepton-pair)

E605, E866 (d/p ratio)

Drell-Yan – W and Z

CDF (W-lepton-asymmetry)

Direct Photon Production

WA70, UA6, E706, ISR, Ua2, CDF, D0

Inclusive Jet Production

CDF, D0

Lepto-production of Heavy Quark

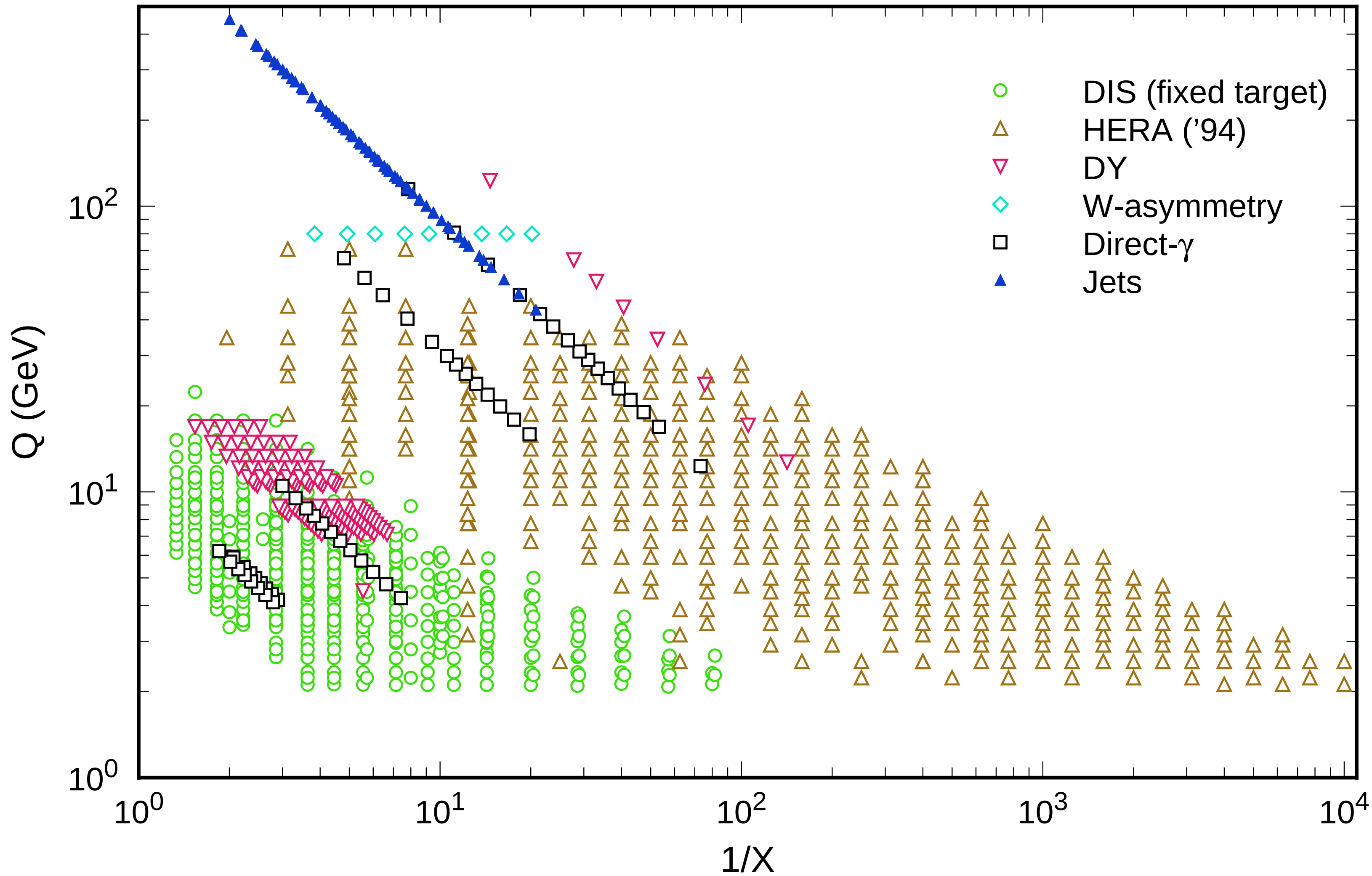
H1, ZEUS

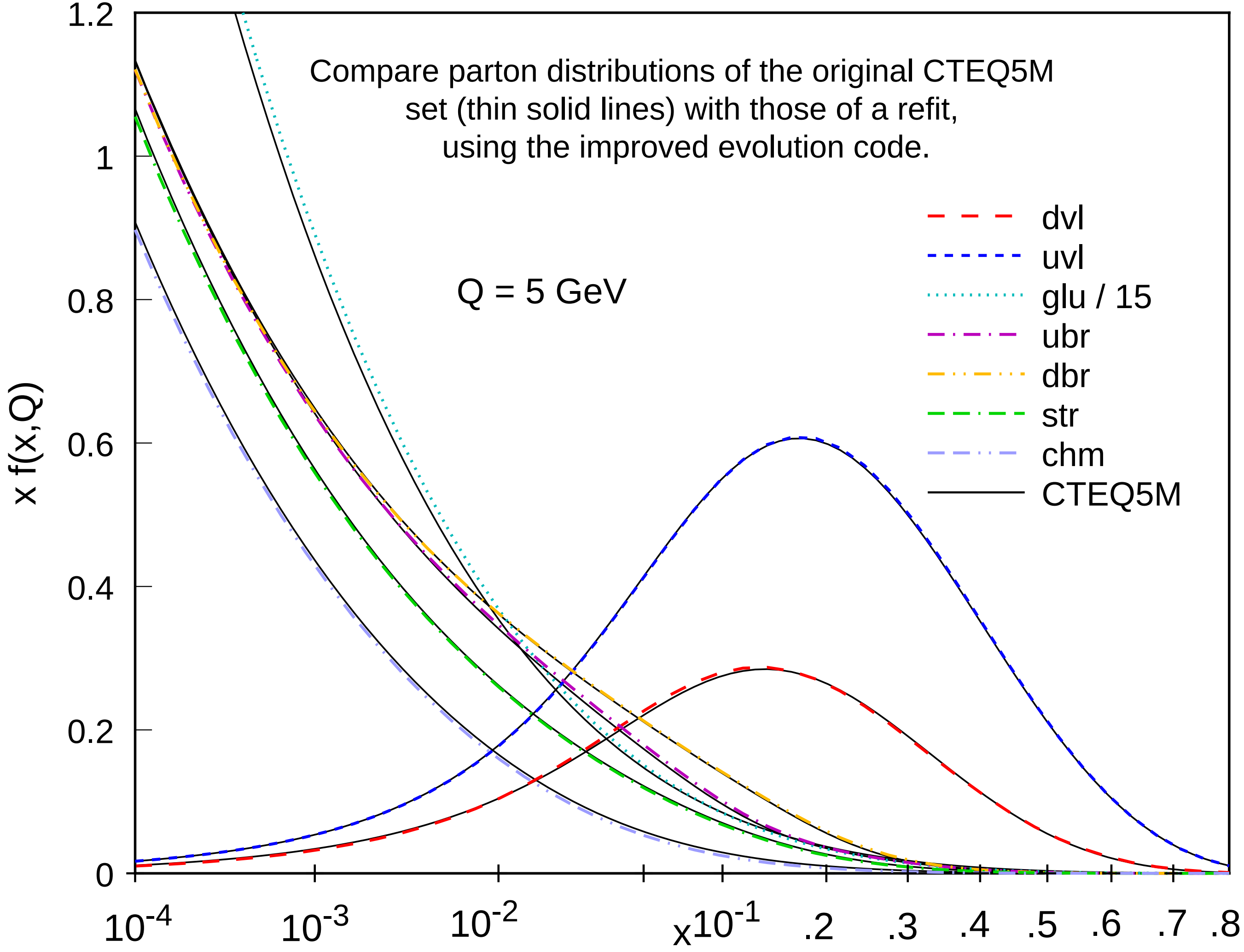
Hadro-production of Heavy Quark

CDF, D0

* Red color indicates "New" for current analysis
Green indicates "often not used" in global analysis

Coverage of Current Experiments in the (x, Q) plane





Compare parton distributions of the original CTEQ5M
set (thin solid lines) with those of a refit,
using the improved evolution code.
Q = 5 GeV
dvl
uvl
glu / 15
ubr
dbr
str
chm
CTEQ5M
x f(x,Q)
x
1.2
1
0.8
0.6
0.4
0.2
0
10^-4
10^-3
10^-2
10^-1
.2
.3
.4
.5
.6
.7
.8

Uncertainties of Physical Quantities X due to Parton Distributions

–Direct approach: the Lagrange multiplier Method

Start with CTEQ global analysis, using χ^2 minimization with a suitably defined χ_{global}^2 function;

--e.g. use CTEQ5M1 as the "best fit".

Add a Lagrange multiplier term proportional to the X-sec. to the χ^2 function and minimize the sum under the constraints of all existing expts;

$$\Psi_\lambda (a_l) = \chi_{global}^2 + \lambda X$$

Explore how χ_{global}^2 (for the complete original data sets) varies as a function of X;

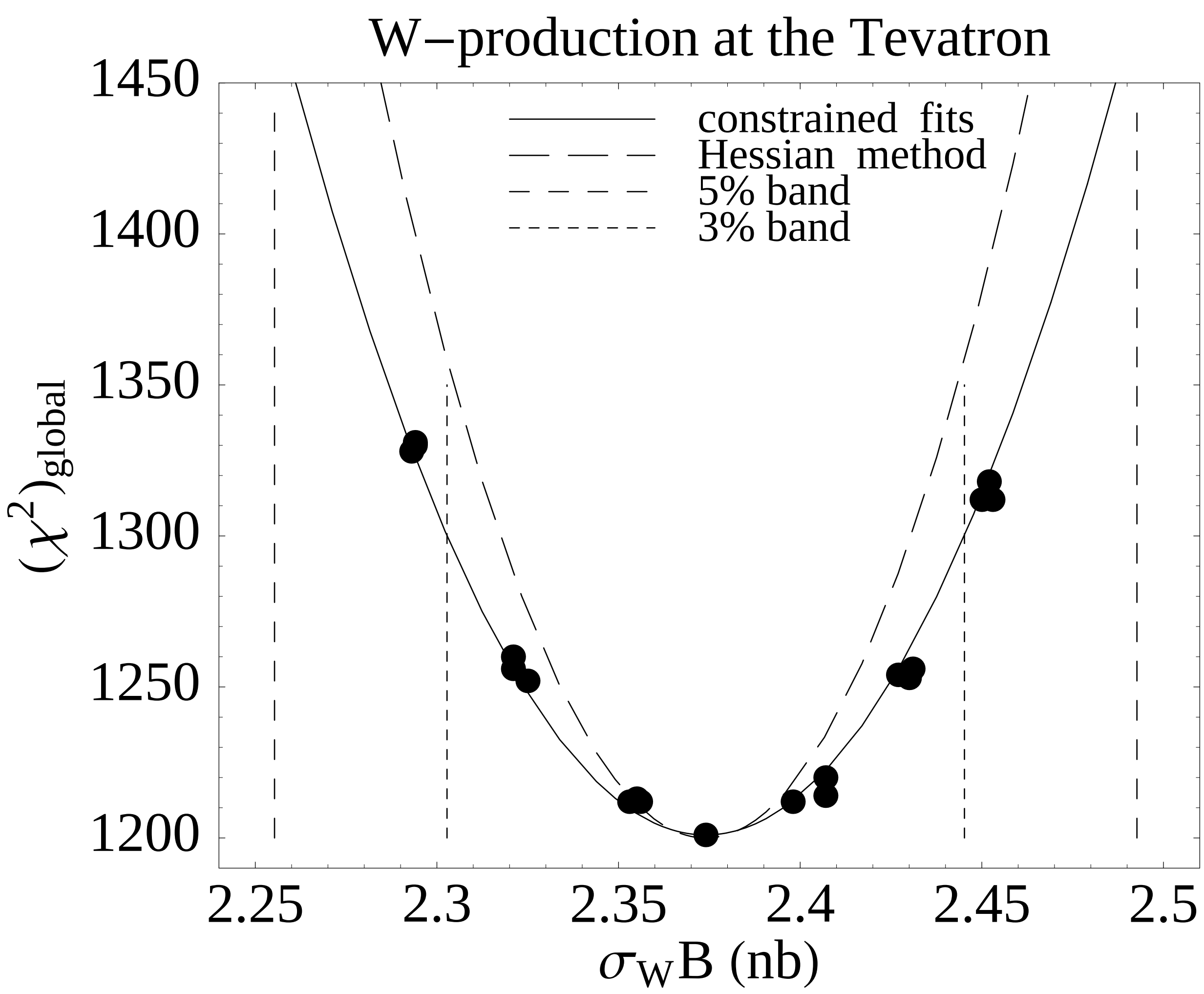
$$\text{minimize } \Psi_\lambda (a_l) \Rightarrow \chi_{global}^2 (X)$$

which takes into account all d.o.f. of PDF parameter space -- cf. plot of χ_{global}^2 vs. σ_W

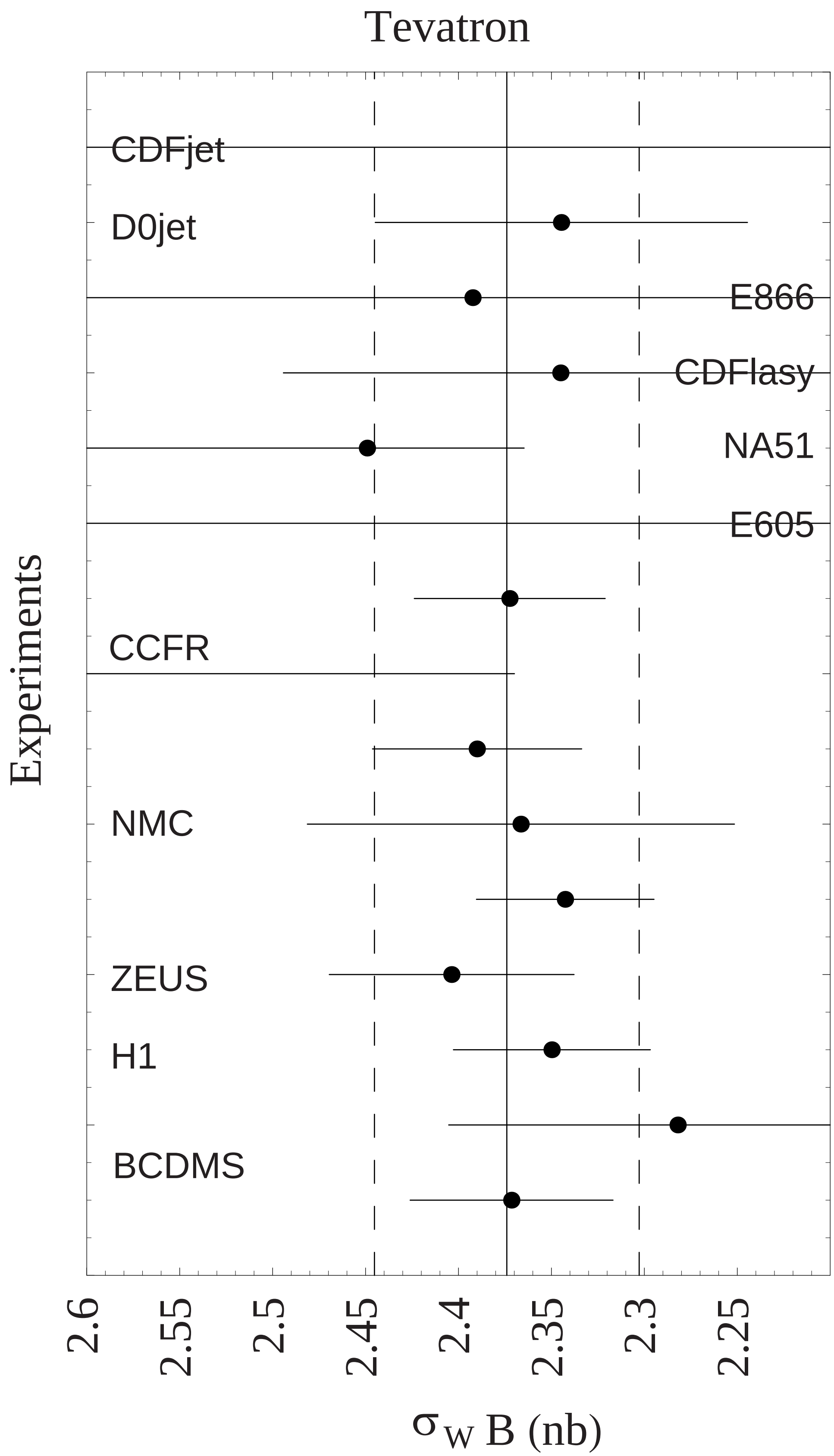
Analyze the probabilities for the individual expts, including systematic errors, if available, and estimate a 68% "one σ" for each expt. n: σ_n;

Combine the errors $\{\sigma_n\}$ of the full set of expts used in the global analysis, and obtain a nominal one σ deviation for X, σ_X.

Example: X = W prod. Xsec. σ_W



W−production at the Tevatron
constrained fits
Hessian method
5% band
3% band
1450
1400
1350
1300
1250
1200
$(\chi^2)_{\rm global}$
2.25
2.3
2.35
2.4
2.45
2.5
$\sigma_W B$ (nb)


Tevatron
CDFjet
D0jet
E866
CDFlasy
NA51
E605
CCFR
NMC
ZEUS
H1
BCDMS
Experiments
2.6
2.55
2.5
2.45
2.4
2.35
2.3
2.25
σ_W B (nb)

Treatment of Correlated Experimental Errors:

Correlated systematic errors a_{jk}
where $k = 1 \ldots n_s$

"True" statistical χ^2:

$$\chi^2 = \sum_j \frac{\left(d_j - t_j\right)^2}{\sigma_j^2} - \sum_{kk'} B_k \left(A^{-1}\right)_{kk'} B_{k'}. \quad (2)$$

The index j labels the data points. The indices k and k' label the source of systematic error and run from 1 to n_s.

B_k is the vector

$$B_k = \sum_j \frac{\left(d_j - t_j\right) a_{jk}}{\sigma_j^2}, \quad (3)$$

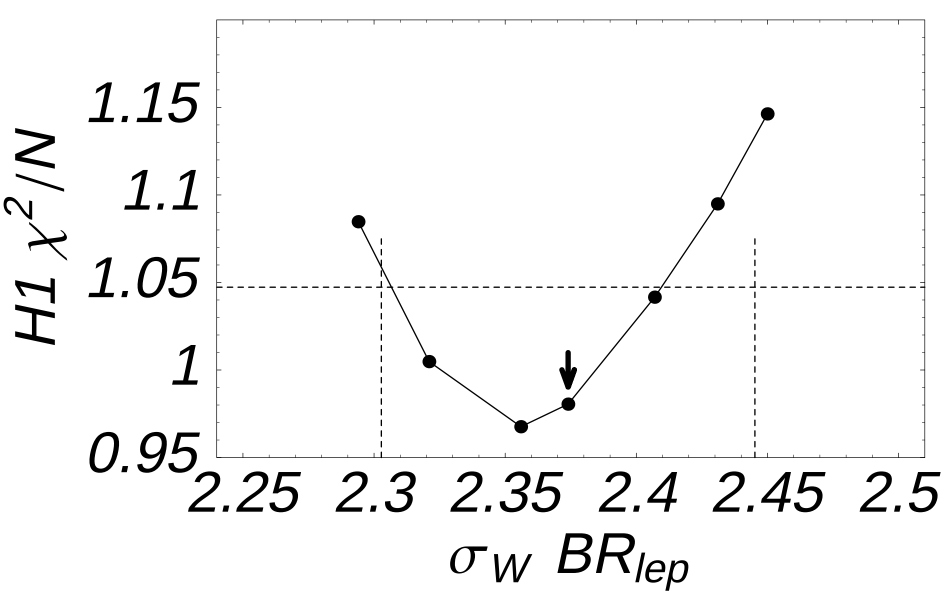
and $A_{kk'}$ is the matrix

$$A_{kk'} = \delta_{kk'} + \sum_j \frac{a_{jk} a_{jk'}}{\sigma_j^2}. \quad (4)$$

Application to the H1 data on F_2

Lagrange multiplier	$\sigma_W \cdot B$ in nb	$\chi^2/172$	probability
3000	2.294	1.0847	0.212
2000	2.321	1.0048	0.468
1000	2.356	0.9676	0.605
0	2.374	0.9805	0.558
-1000	2.407	1.0416	0.339
-2000	2.431	1.0949	0.187
-3000	2.450	1.1463	0.092

χ^2/N of the H1 data, including error correlations, compared to PDFs obtained by the Lagrange multiplier method for constrained values of σ_W



The Standard Error-Matrix (Hessian) Method

At the minimum of the χ^2 function,

$$\chi^2 = \chi_0^2 + \frac{1}{2}\sum_{i,j} H_{ij} y_i y_j$$

where $y_i = a_i - a_{0i}$ is the displacement from the minimum, and H_{ij} is the *Hessian*.

Once H_{ij} is determined from a global fit,

* the eigenvalues and complete orthonormal set of eigenvectors of H_{ij} give insight on the uncertainties of PDF's around the minimum;

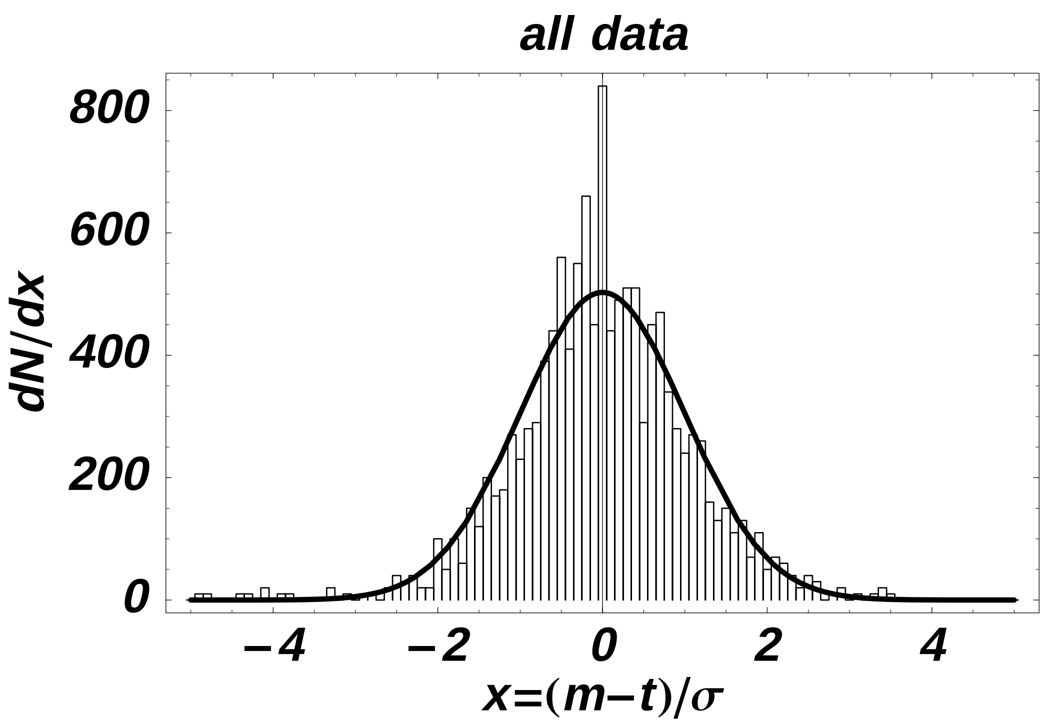
* allow to identify systematically the particular degrees of freedom which need further experimental input in future global analyses;

* allow a uniform way to determine uncertainties of all physical variables (X) of interest:

$$\Delta X = \sqrt{2\sum_{i,j} \frac{\partial X}{\partial a_i}(H^{-1})^{ij}\frac{\partial X}{\partial a_j}}.$$

When applied to study the uncertainty of the W cross-section at hadron colliders, the results are in agreement with those of using the Lagrange multiplier method.

Comparison of data and CTEQ5m ("theory")



The histogram includes all data used in the fit except jet production.

The curve has no adjustable parameters; it's just

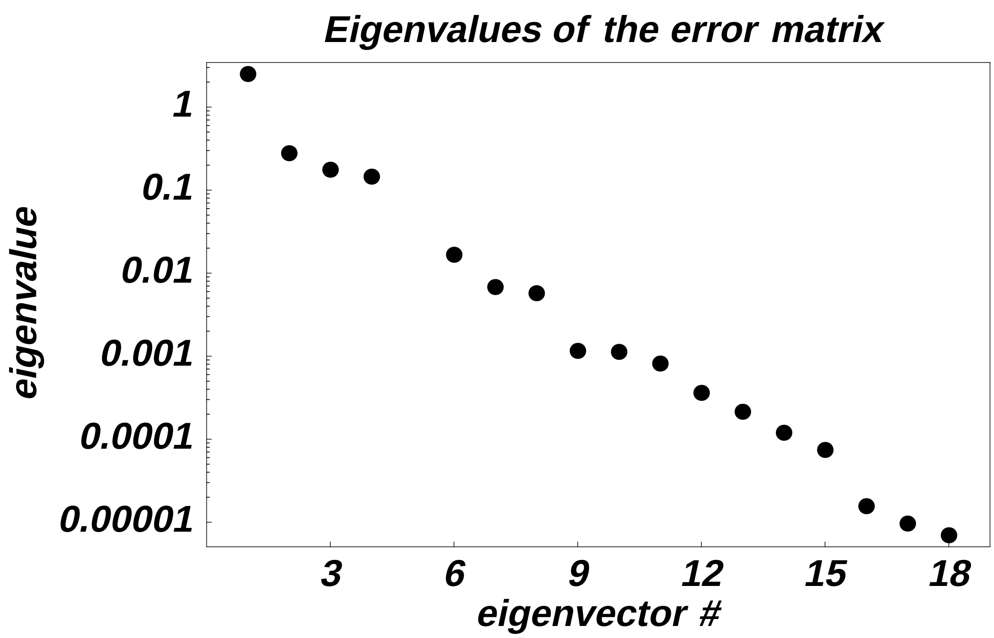
$$\mathcal{N}\,\exp(-x^2)/\sqrt{2\pi}$$

where $\mathcal{N}$ is the number of data points. The area under the curve (or histogram) is $\mathcal{N}$.

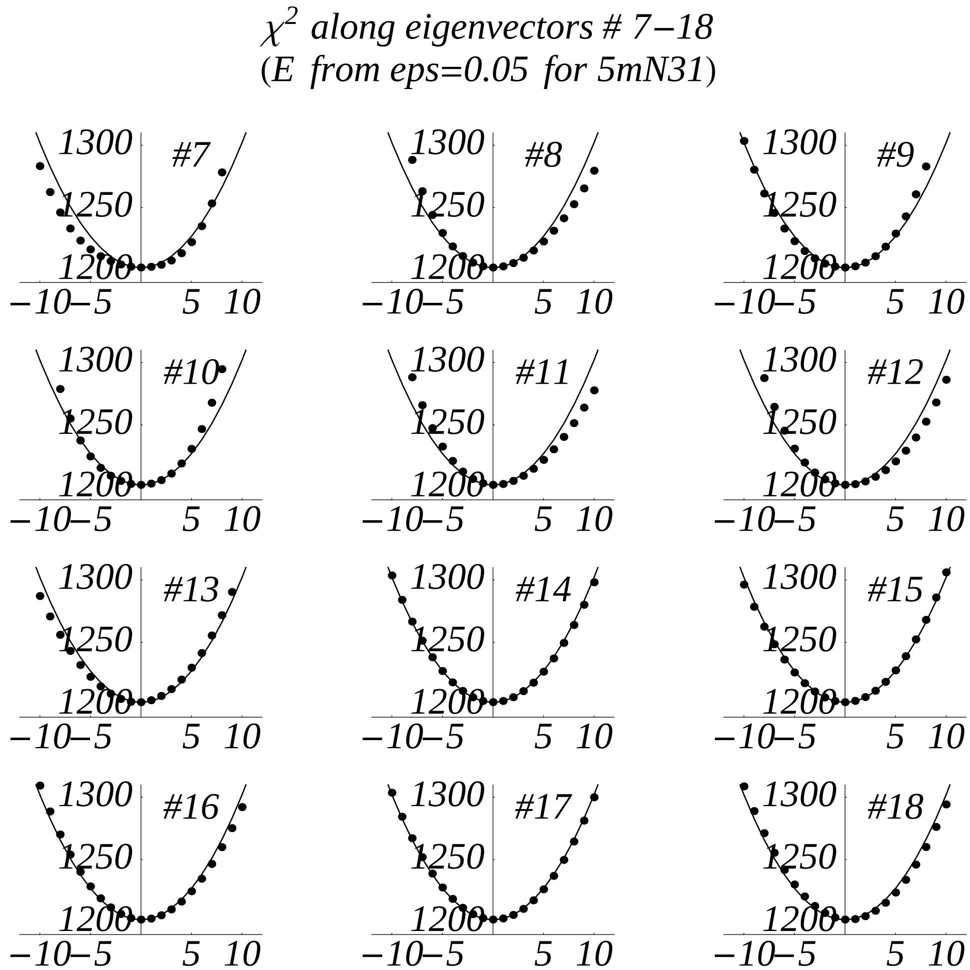
Differences $m_i - t_i$ are within the published measurement errors.

At least *globally* the distribution of fluctuations is Gaussian with the right width.

Eigenvalues of the error matrix



Variation of χ^2 along the eigenvectors



The parabola has no free parameters, and is the same on each graph $= 1200 + x^2$.

The variable x is $C/\sqrt{e}$, where the displacement from the minimum is $C\,U$.
(U=eigenvector, e=eigenvalue).

The minimum is approximately quadratic, at least out to $\Delta\chi^2 = 100$. (Compare $\chi^2_{\text{min}} = 1200$.)

These eigenvectors of the error matrix have the smallest eigenvalues – *i.e.*, the steep directions.

Results of the MSU Group on Uncertainties of PDFs and Phys. Observables

(based on apparatus of CTEQ global analysis)

Lagrange Multiplier Method
– for specific physical variables, say X –

- Uncertainty of X: ΔX – within "one s.d." allowed by global fit to all pertinent data sets;
- Three sets of PDFs, with $\Delta X = -1, 0, 1$, which can be used for relevant applications;
- This can be done for any X: e.g.
  · $\sigma_{W,Z}$, moments of y and p_T distributions;
  · various Parton Luminosities at a given $\sqrt{s}$;
  · σ_{Higgs}, $\cdots$

Hessian Method for Parton Distributions and for general physical variables

- Eigenvalues and Eigenvectors of the Hessian;
- Given physical variable X and the vector $\partial X/\partial a_i$, can calculate $\Delta X = \sqrt{2\sum_{i,j} \frac{\partial X}{\partial a_i}(H^{-1})_{ij}\frac{\partial X}{\partial a_j}}$;
- Systematically identify "Flat directions" where parametrization can be tightened, or improvements in experimental constraints are needed;
- Collection of PDFs on the surface of "one s.d." of the Hessian, suitable for probing uncertainties of a variety of physical variables.

Some Details of
Strange Sea Analysis of Barone etal.

- Data sets*
  CC-DIS $(\nu, \bar{\nu})$: BEBC, CDHS, CDHSW
  NC-DIS (e, μ): BCDMS, NMC, H1
  Drell-Yan: E605, NA51, E866
  (* The CCFR data, coming from a different preanalysis, are not included.)
- Data have been properly reanalyzed: Bin center corrections, EW radiative corrections, corrections for nuclear and isoscalarity effects are applied.
- Error correlations have been taken into account.
- Use the Fixed (3) Flavor Number Scheme to treat charm mass.
- The kinematic cuts: $Q^2 \geq 3.5$ GeV^2, $W^2 \geq 10$ GeV^2 (to avoid higher-twist effects). For the CDHSW data, exclude controversial $x < 0.1$ region.

The χ^2's of three fits

# pts	χ^2 `fit1` ($\kappa = 0.67$)	χ^2 `fit1b` ($\kappa = 0.5$)	χ^2 `fit2` ($s(x) \neq \bar{s}(x)$)
2657	2430.8	2492.4	2405.0