Hessian PDF reweighting meets the Bayesian methods

Hannu Paukkunen

University of Jyväskylä, Finland
Helsinki Institute of Physics, Finland

based on an article written in collaboration with

Pia Zurita

Reference: arXiv:1402.6623 [hep-ph]

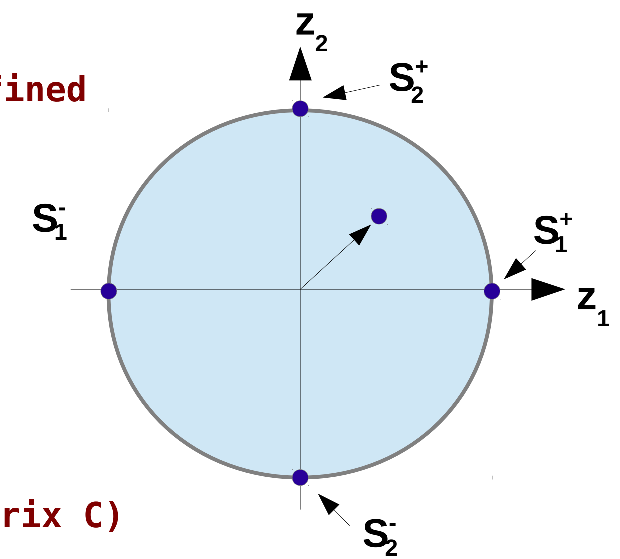
The Hessian reweighting

- **Builds on the standard Hessian method to quantify PDF errors**

$$\chi^2\{a\} = \sum_k \left[\frac{X_k^{\rm theory}[f] - X_k^{\rm data}}{\delta_k^{\rm data}} \right]^2 \approx \chi_0^2 + \sum_{ij} \delta a_i H_{ij} \delta a_j \approx \chi_0^2 + \sum_i z_i^2$$

- **In the case of a global tolerance, the error sets are defined in the z space**

$$\begin{aligned} z(S_0) &= (0,0,...,0)\,, \\ z(S_1^\pm) &= \pm\sqrt{\Delta\chi^2}\,(1,0,...,0) \\ z(S_2^\pm) &= \pm\sqrt{\Delta\chi^2}\,(0,1,...,0) \\ &\vdots \\ z(S_{N_{\rm eig}}^\pm) &= \pm\sqrt{\Delta\chi^2}\,(0,0,...,1) \end{aligned}$$



- **Add the contribution of new data {y} (with covariance matrix C) to the expression above**

$$\chi^2_{\rm new} \equiv \chi_0^2 + \sum_k^{N_{\rm eig}} z_k^2 + \sum_{i,j=1}^{N_{\rm data}} \left(y_i[f] - y_i\right) C_{ij}^{-1} \left(y_j[f] - y_j\right)$$

and estimate the theory values $y_i\,[f]$ by

$$D_{ik} \equiv \frac{y_i\left[S_k^+\right] - y_i\left[S_k^-\right]}{2}$$

$$y_i\,[f] \approx y_i\,[S_0] + \sum_{k=1}^{N_{\rm eig}} \frac{\partial y_i\,[S]}{\partial z_k}\bigg|_{S=S_0} z_k \approx y_i\,[S_0] + \sum_{k=1}^{N_{\rm eig}} D_{ik} w_k$$

$$w_k \equiv \frac{z_k}{\sqrt{\Delta\chi^2}}$$

The Hessian reweighting

- The new global minimum is obtained by the matrix equation

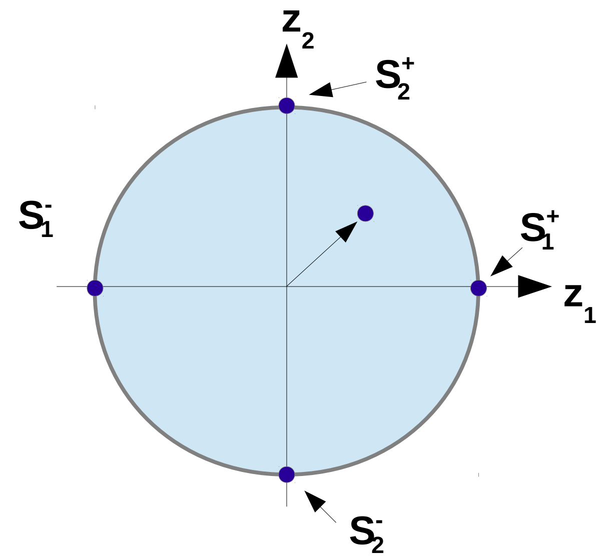
$$B_{kn} = \sum_{i,j} D_{ik} C^{-1}_{ij} D_{jn} + \Delta\chi^2 \delta_{kn}$$

$$\vec{\mathbf{w}}^{\rm min} = -\mathbf{B}^{-1}\vec{\mathbf{a}}$$

$$a_k = \sum_{i,j} D_{ik} C^{-1}_{ij} \left(y_j\left[S_0\right] - y_j\right)$$

- The corresponding set of PDF is given by

$$f^{\rm new} \approx f_{S_0} + \sum_{k=1}^{N_{\rm eig}} \left(\frac{f_{S_k^+} - f_{S_k^-}}{2}\right) w_k^{\rm min}$$



- The new χ^2 can be written as

$$\chi^2_{\rm new} = \chi^2_{\rm new}\Big|_{\vec{\mathbf{w}}=\vec{\mathbf{w}}^{\rm min}} + \sum_{ij} \delta w_i B_{ij} \delta w_j$$

...and the new PDF error sets defined by diagonalizing the new "Hessian matrix" B

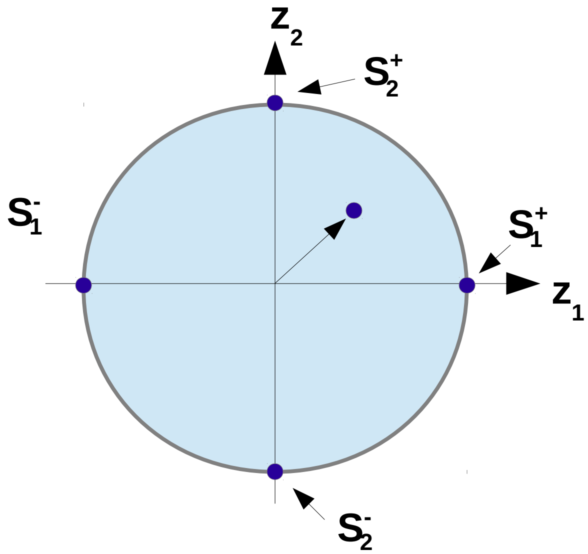
- The result is a new central set of PDFs + error sets

- Need to evaluate the observables once with the central and error sets

The Hessian reweighting

- An important characteristic is the “penalty” term

$$P \equiv \Delta\chi^2 \sum_{k=1}^{N_{\rm eig}} (w_k^{\rm min})^2$$ = the growth of the original χ^2



$P \ll \Delta\chi^2 \longrightarrow$ the new data compatible with the original PDFs

$P \gtrsim \Delta\chi^2 \longrightarrow$ tension with the original PDFs

- The new PDFs satisfy all the relevant sum rules e.g.

$$\int_0^1 dx x \sum_f f^{\rm new} = \int_0^1 dx x \sum_f f_{S_0} + \sum_k \frac{w_k^{\rm min}}{2} \left[\int_0^1 dx x \sum_f f_{S_k^+} - \int_0^1 dx x \sum_f f_{S_k^-} \right]$$
$$= 1 + \sum_k \frac{w_k^{\rm min}}{2} [1-1] = 1.$$

- The new PDFs also satisfy the DGLAP equation (since DGLAP is a linear equation)

- The new PDFs can be used consistently in any perturbative calculations

Bayesian methods

- Construct PDF replicas from the Hessian error sets by... [*JHEP 1208 (2012) 052*]

$$f_k \equiv f_{S_0} + \sum_i^{N_{\mathrm{eig}}} \left(\frac{f_{S_i^+} - f_{S_i^-}}{2} \right) R_{ik} \qquad \langle \mathcal{O} \rangle = \frac{1}{N_{\mathrm{rep}}} \sum_{k=1}^{N_{\mathrm{rep}}} \mathcal{O}\left[f_k\right]$$

Gaussian random numbers

Observable computed with replica k

...and use the Bayesian methods to reweight PDFs

$$\langle \mathcal{O} \rangle_{\mathrm{new}} = \frac{1}{N_{\mathrm{rep}}} \sum_{k=1}^{N_{\mathrm{rep}}} \omega_k \, \mathcal{O}\left[f_k\right]$$

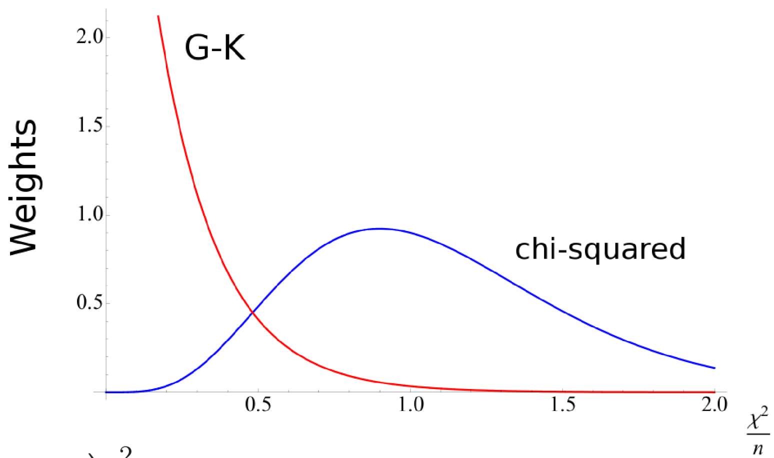
- NNPDF collaboration:
  [Nucl.Phys. B849 (2011) 112-143, Nucl.Phys. B855 (2012) 608-638]

$$\chi_k^2 = \sum_{i,j=1}^{N_{\mathrm{data}}} \left(y_i[f_k] - y_i\right) C_{ij}^{-1} \left(y_j[f_k] - y_j\right)$$

$$\omega_k^{\mathrm{chi-squared}} = \frac{\left(\chi_k^2\right)^{(N_{\mathrm{data}}-1)/2} \exp\left[-\chi_k^2/2\right]}{(1/N_{\mathrm{rep}}) \sum_{k=1}^{N_{\mathrm{rep}}} \left(\chi_k^2\right)^{(N_{\mathrm{data}}-1)/2} \exp\left[-\chi_k^2/2\right]}$$

Giele & Keller: [Phys.Rev. D58 (1998) 094023]

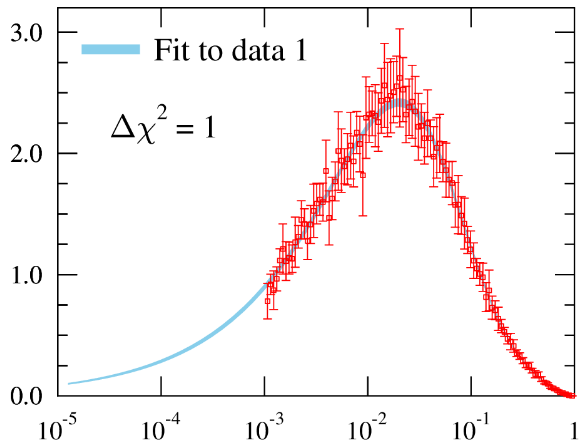
$$\omega_k^{\mathrm{GK}} = \frac{\exp\left[-\chi_k^2/2\right]}{(1/N_{\mathrm{rep}}) \sum_{k=1}^{N_{\mathrm{rep}}} \exp\left[-\chi_k^2/2\right]}$$

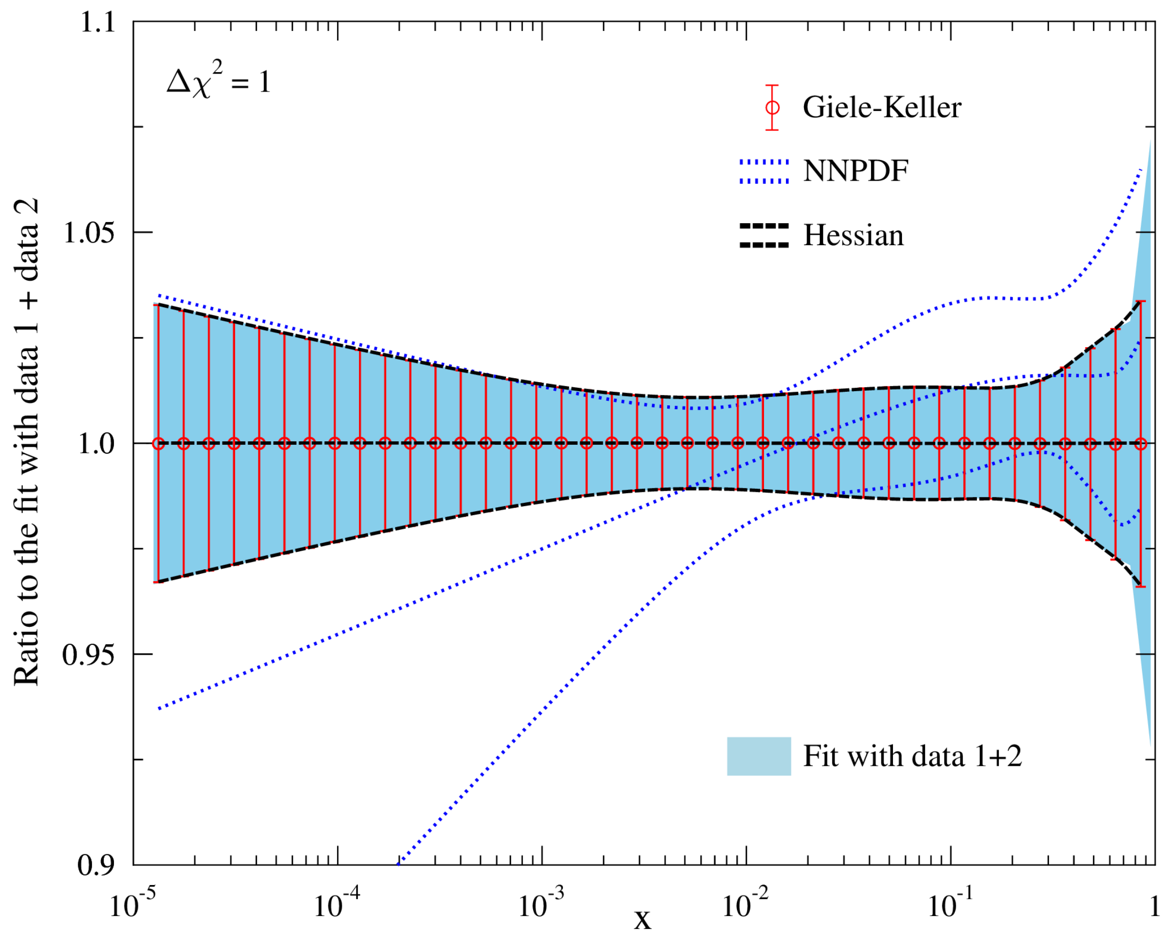


- The reweighting penalty becomes $P = \Delta\chi^2 \sum_i^{N_{\mathrm{eig}}} \left(\frac{1}{N_{\mathrm{rep}}} \sum_k^{N_{\mathrm{rep}}} \omega_k R_{ik} \right)^2$

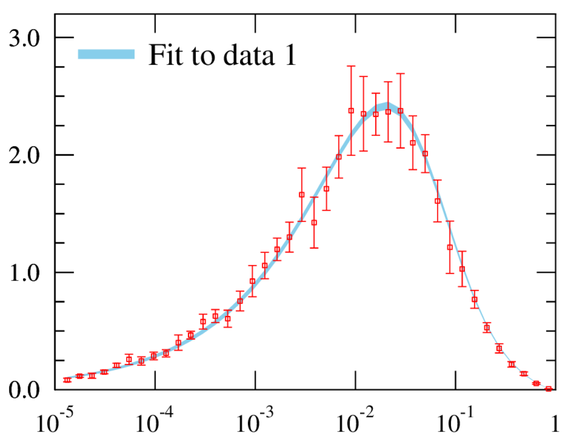
Simple example $\Delta\chi^2$=1

- **Construct a set of data for** $g(x) = a_0 x^{a_1}(1-x)^{a_2} e^{x a_3}(1+x e^{a_4})^{a_5}$
- **Fit and construct the Hessian error sets using $\Delta\chi^2$=1**





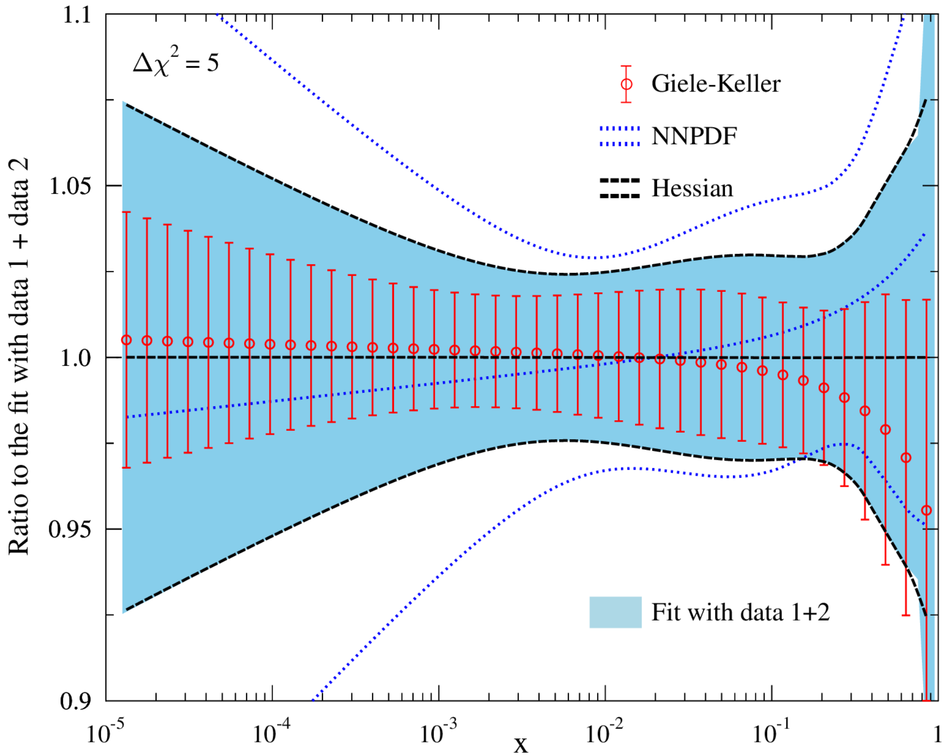
- **Apply the reweighting methods on a second set of data and compare to the fit including <u>both</u> data sets**



- **Hessian reweighting, Bayesian reweighting with GK weights, and direct fit agree**
- **Bayesian reweighting with NNPDF weights differs from the rest**

Simple example $\Delta\chi^2=5$

- **Same example but with non-zero tolerance $\Delta\chi^2=5$**



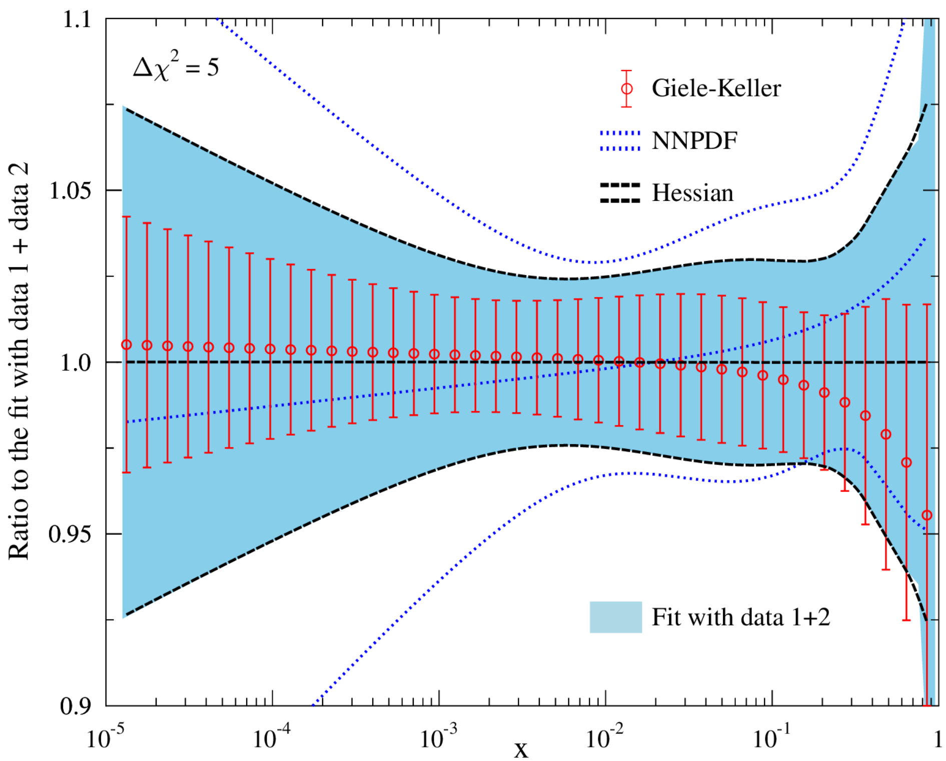
- **Rescale the values of χ^2**

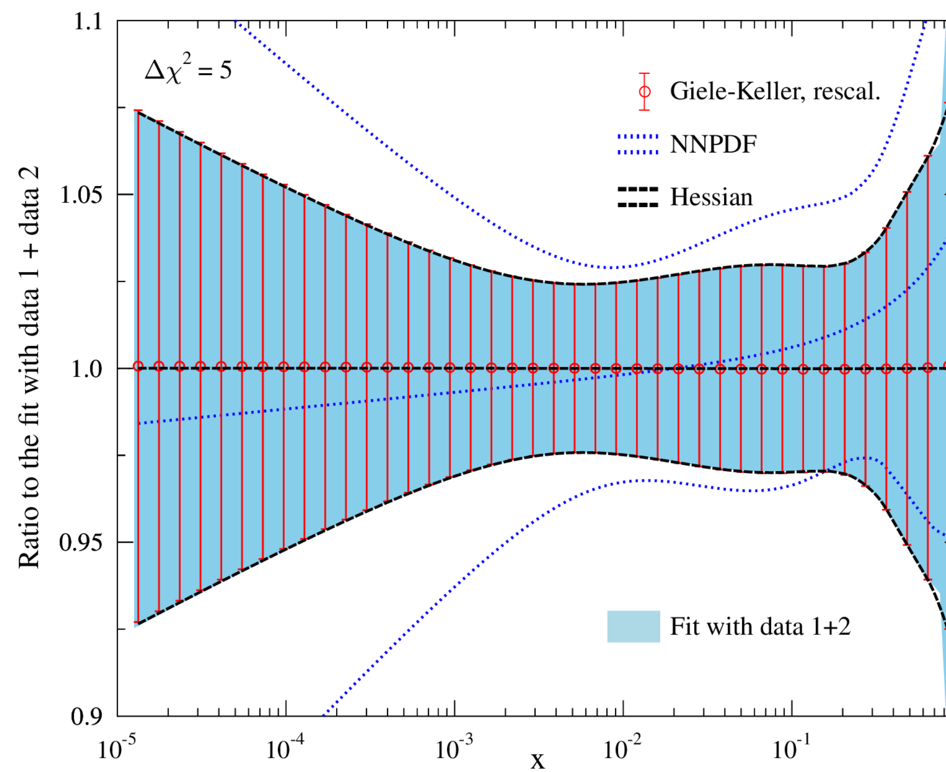
$$\chi_k^2 \longrightarrow \frac{\chi_k^2}{\Delta\chi^2}$$

when computing the Giele-Keller weights

Simple example $\Delta\chi^2=5$

- **Same example but with non-zero tolerance $\Delta\chi^2=5$**





- **Rescale the values of χ^2**

$$\chi_k^2 \longrightarrow \frac{\chi_k^2}{\Delta\chi^2}$$

when computing the Giele-Keller weights

Agreement restored!

The equivalence of the Hessian and rescaled Giele-Keller reweighting can also be shown mathematically

Application: CMS inclusive jets

- Reweight CTEQ6.6 ($\Delta\chi^2$=100) with CMS 7TeV inclusive jets [Phys.Rev. D87 (2013) 112002]

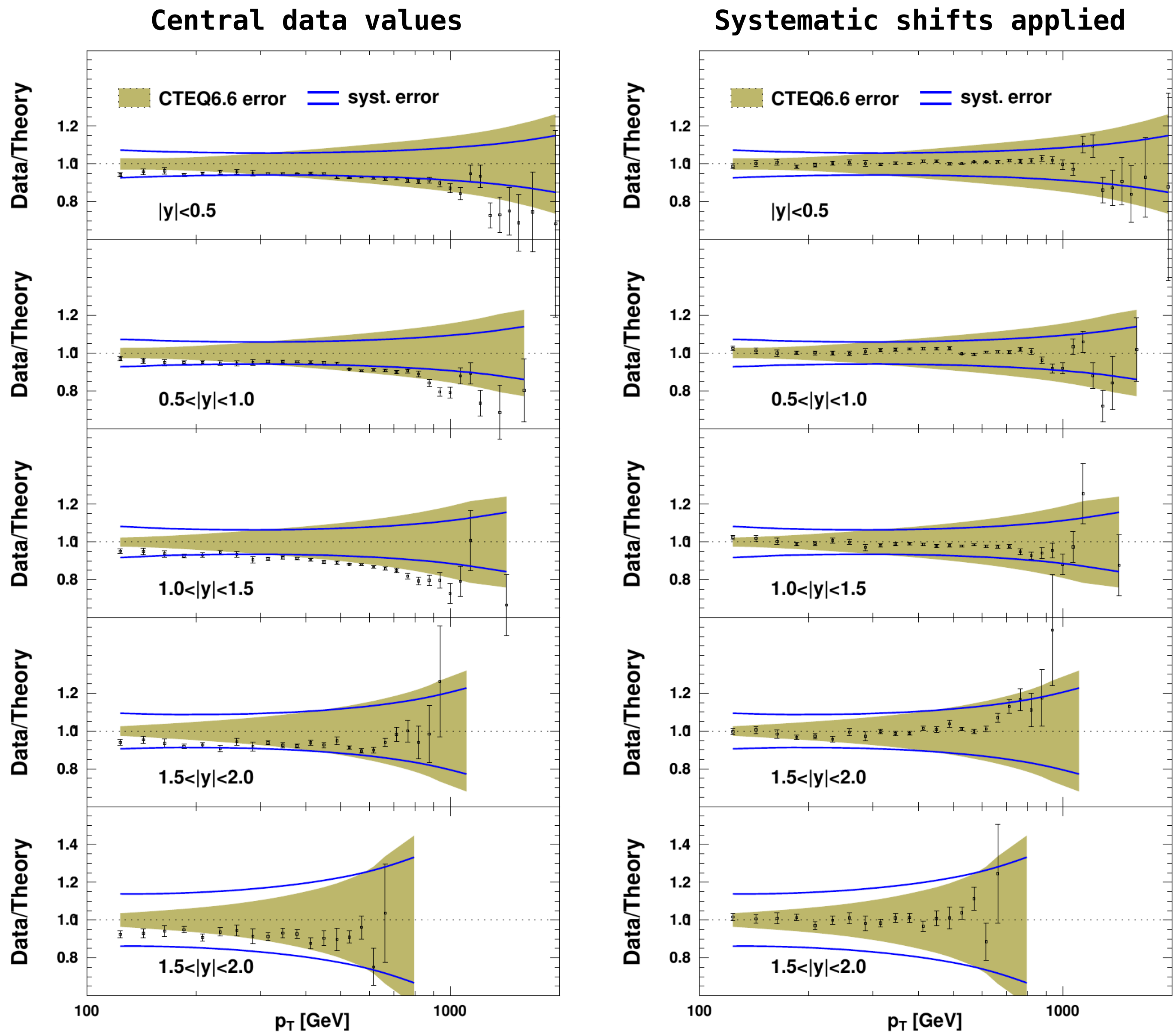
- Computations with FASTNLO (μ=pT/2, $\alpha_s(M_Z)$=0.118)

- Before the reweighting CTEQ6.6 overshoots the data by some 5%

- Can be largely "hidden" to the correlated systematic errors

- χ^2/N=2.1

Fairly large, indicates that these data should have an impact!



Application: CMS inclusive jets

- **Reweighted gluon PDF**

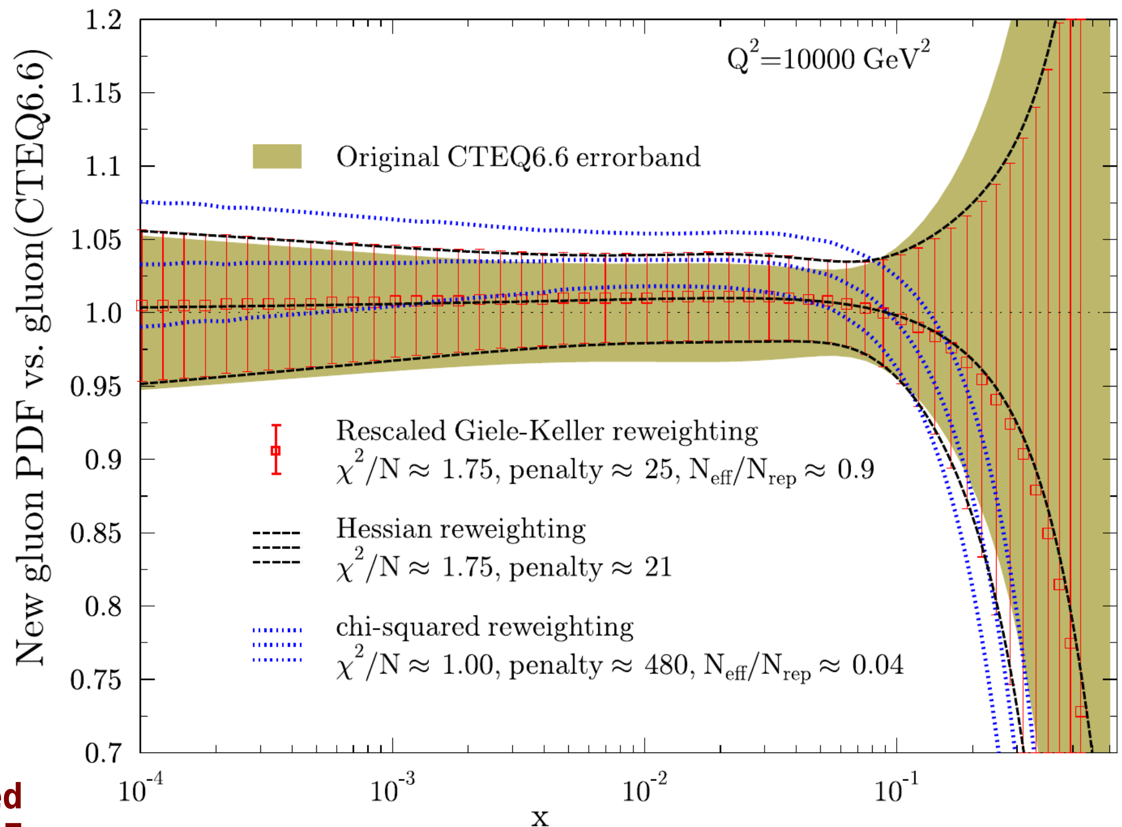
- **Hessian reweighting and rescaled GK reweighting:**

  penalty 21 << $\Delta\chi^2$=100

  Could have included these jet data to CTEQ6.6 fit within the $\Delta\chi^2$=100 tolerance

  fairly large χ^2/N=1.75

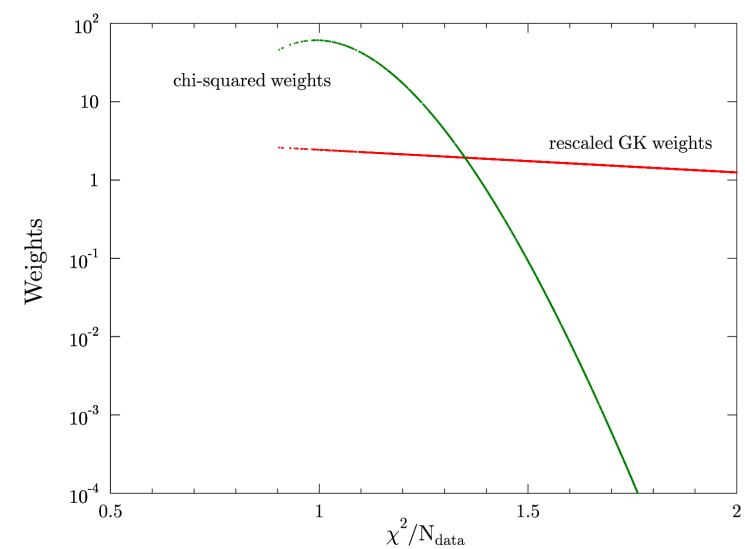
  The Tevatron Run-1 jet data (included in CTEQ6.6) prefer "harder" gluon PDF than these CMS data



- **NNPDF reweighting: penalty 480 >> $\Delta\chi^2$=100, χ^2/N=1.0**

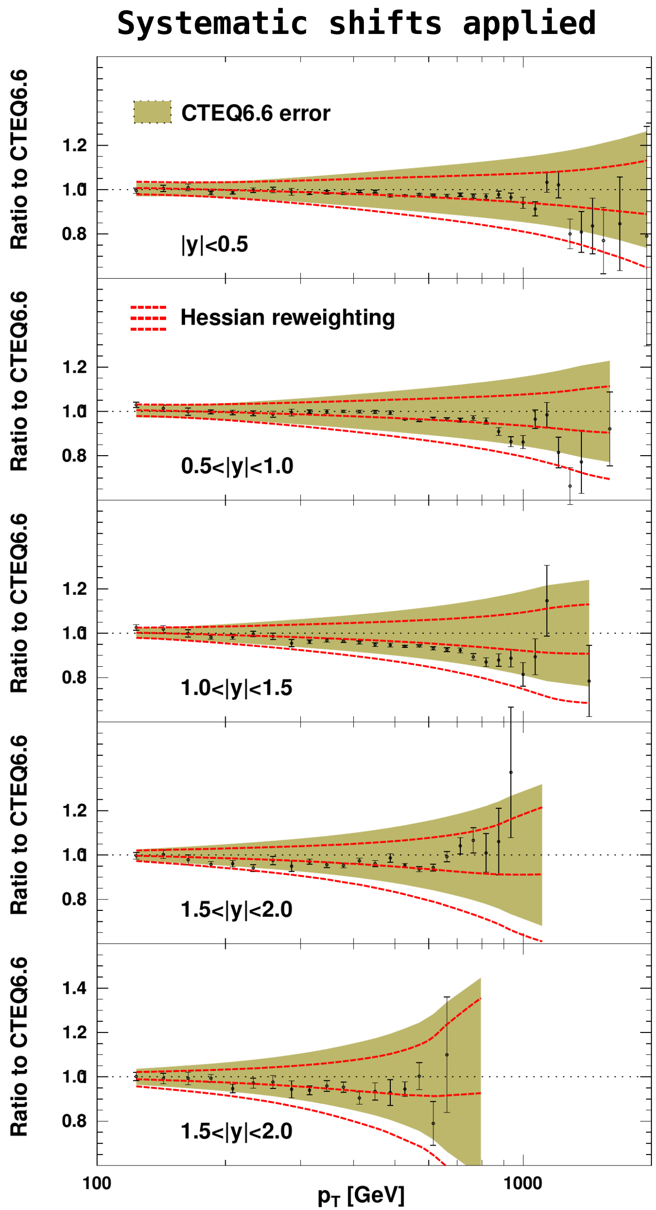
  Practically "ignores" the other data in CTEQ6.6

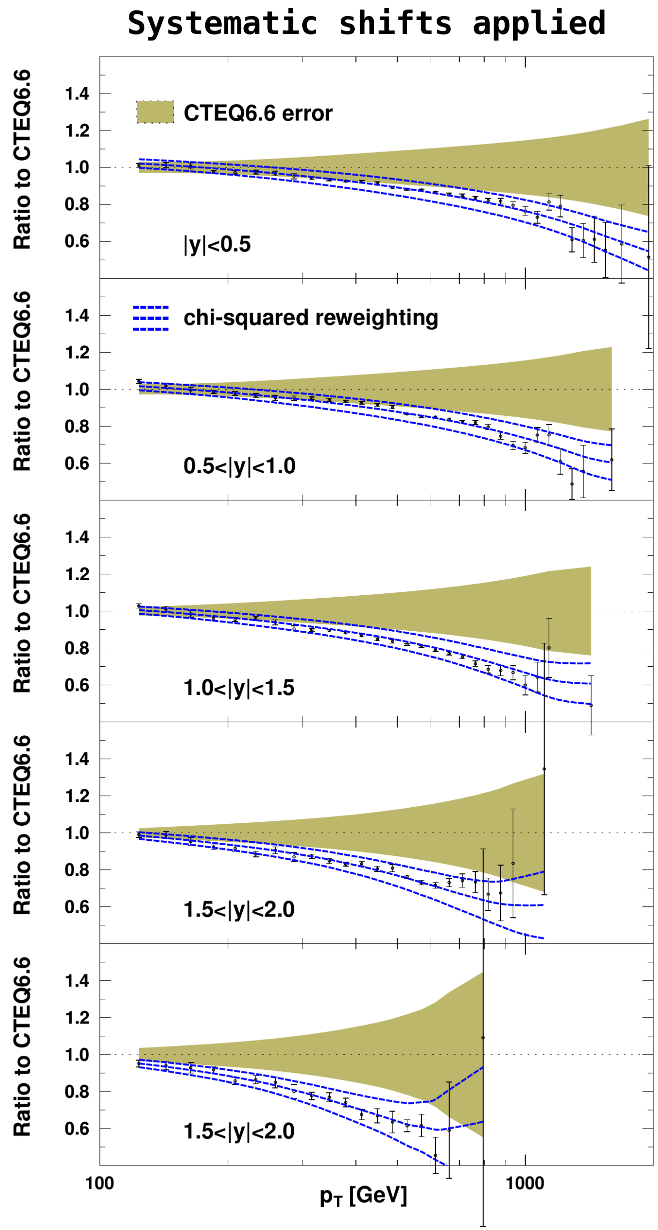
  Based on our simple example, does not correspond to adding these data in the CTEQ6.6 global fit



Application: CMS inclusive jets

- The jet cross sections after the reweighting
- The Hessian method brings the cross sections somewhat below CTEQ6.6
- NNPDF weights cause the predictions to fall clearly downwards
- The systematic shifts depend significantly on the reweighting method





Non-linear extension: MSTW2008

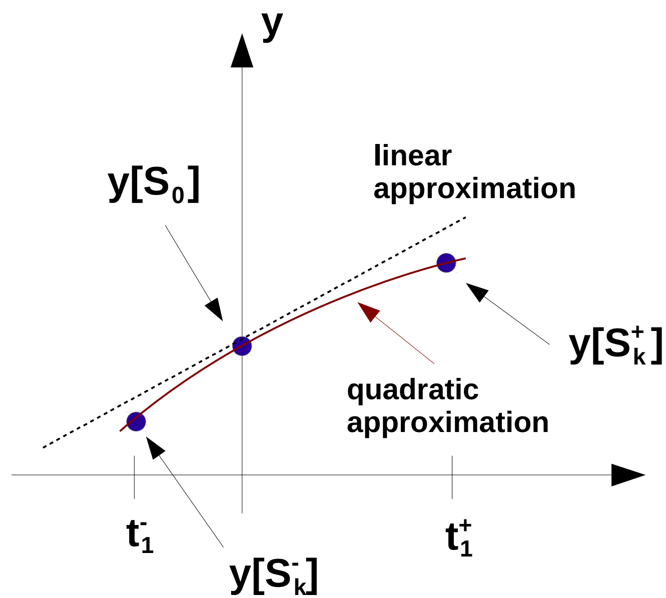
- MSTW2008 uses a "dynamic tolerance" in definining the PDF error sets $z_i(S_k^{\pm}) \equiv \pm t_k^{\pm} \delta_{ik}$
  (20 eigenvectors available, 68% and 90% confidence-level sets)

- Can improve the linear approximations by including also quadratic terms when evaluating the observables

$$y_i\left[S\right] = y_i\left[S_0\right] + \sum_{k=1}^{N_{\rm eig}} \frac{1}{2}\left[\frac{y_i\left[S_k^+\right] - y_i\left[S_0\right]}{t_k^+/t_k^-} - \frac{y_i\left[S_k^-\right] - y_i\left[S_0\right]}{t_k^-/t_k^+}\right] w_k$$

$$+ \sum_{k=1}^{N_{\rm eig}} \frac{t_k^+ + t_k^-}{4}\left[\frac{y_i\left[S_k^+\right] - y_i\left[S_0\right]}{t_k^+} + \frac{y_i\left[S_k^-\right] - y_i\left[S_0\right]}{t_k^-}\right] w_k^2$$

$$w_k \equiv \frac{z_k}{\frac{1}{2}\left(t_k^+ + t_k^-\right)}$$



- Can also improve the quadratic χ^2 profile by

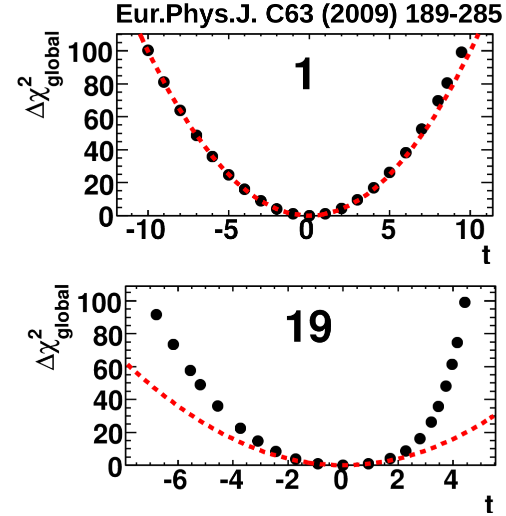
$$\sum_k^{N_{\rm eig}} z_k^2 \longrightarrow \sum_k^{N_{\rm eig}} a_k z_k^2 + b_k z_k^3$$

$$a_k = \frac{t_k^-(T_k^+/t_k^+)^2 - t_k^+(T_k^-/t_k^-)^2}{t_k^+ + t_k^-}$$

$$b_k = \frac{(T_k^+/t_k^+)^2 - (T_k^-/t_k^-)^2}{t_k^+ + t_k^-}.$$

$(T^+)^2 = \Delta\chi^2(t^+)$
$(T^-)^2 = \Delta\chi^2(t^-)$

Eur.Phys.J. C63 (2009) 189-285



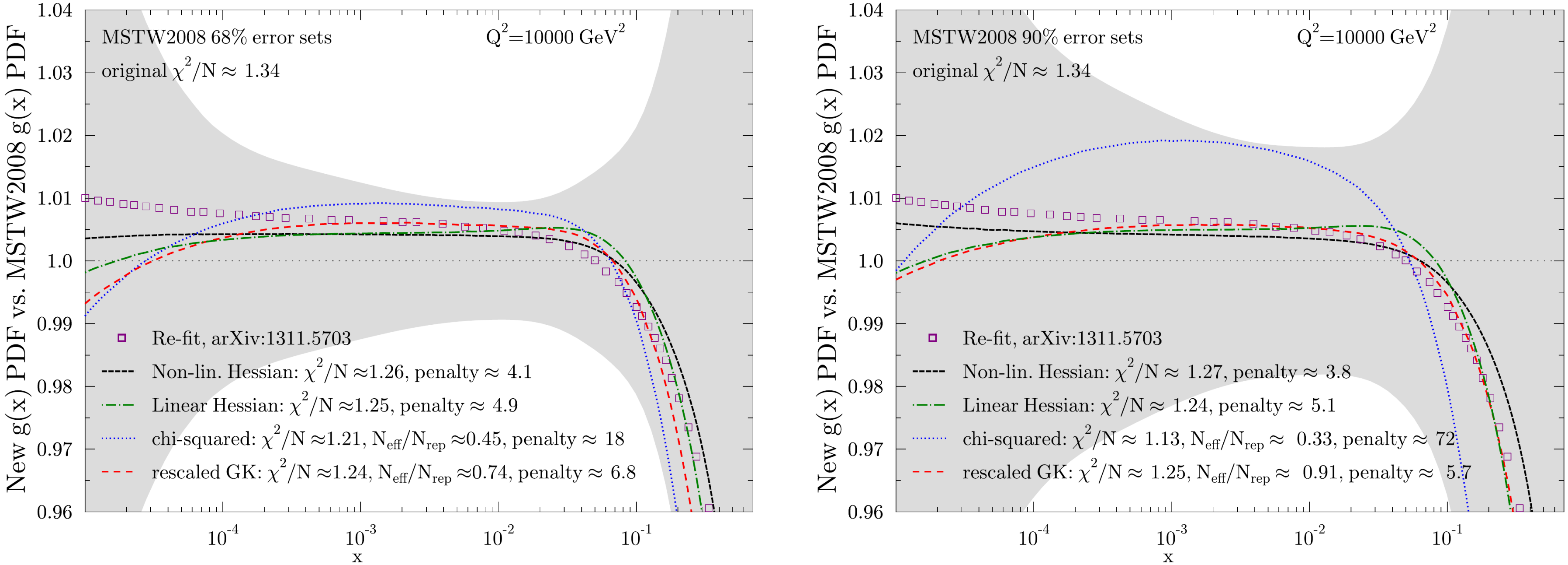
- The χ^2 is no longer quadratic and the minimization has to be done numerically

- Use the same CMS jet data to reweigth MSTW2008

Non-linear extension: MSTW2008

- Compare to the direct fit "MSTWCMS" [arXiv:1311.5703] based on CMS and ATLAS data (28 open parameters + the data normalizations)

- Also linear Hessian reweighting and rescaled Giele-Keller with $\Delta\chi^2(68\%)=10$, $\Delta\chi^2(90\%)=40$



- Hessian and rescaled Giele-Keller reweighting give consistent results – close to the "exact" result

- The NNPDF weights predicts too pronounced effects (clearly larger penalty)

- Non-linearities important when the PDF errors are larger

Non-linear extension: MSTW2008

- Compare the cross sections after the reweighting

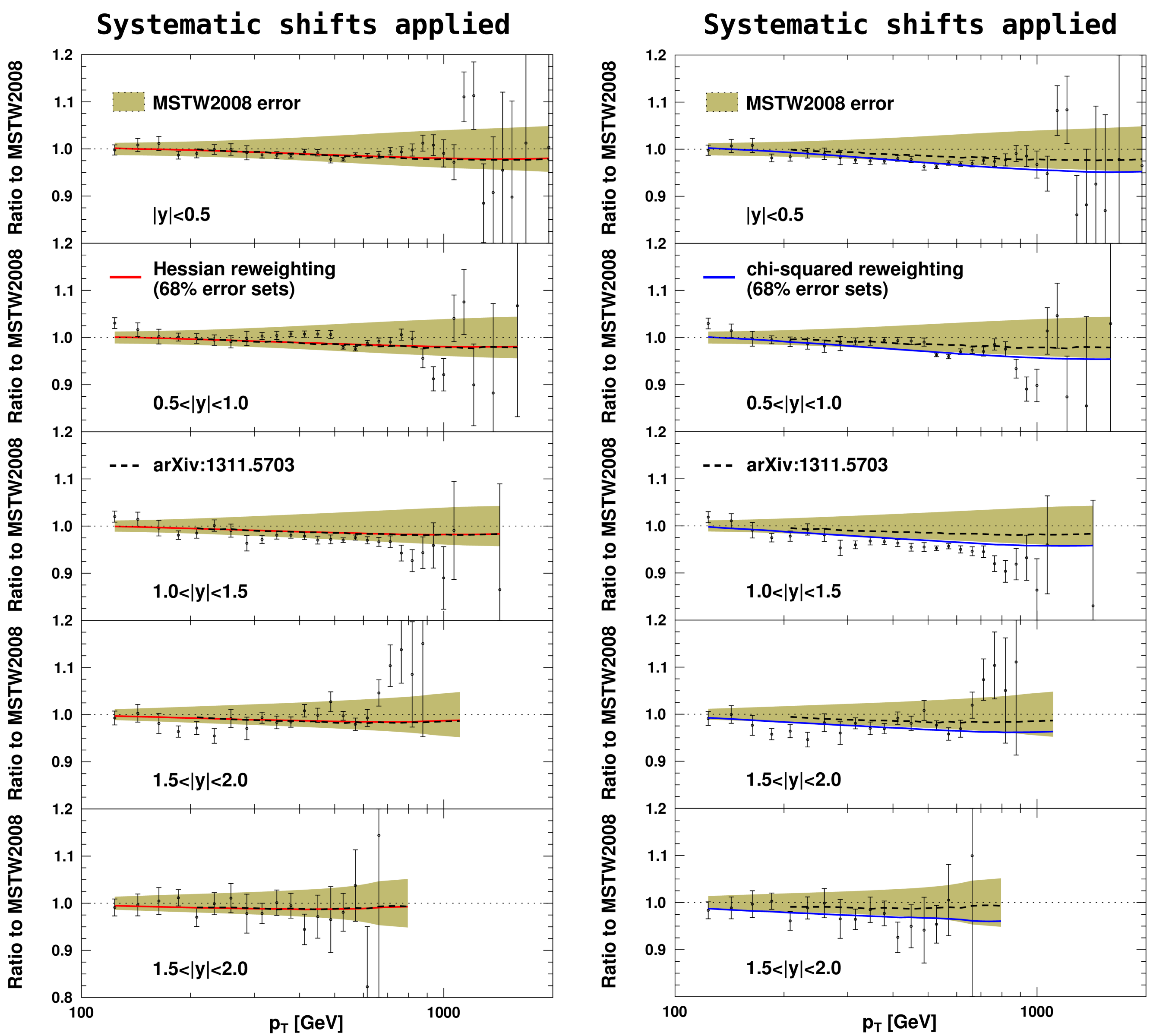
- The 68% error sets

  Hessian reweighting:

  A small downward shift, compatible with MSTWCMS

  Bayesian reweighting with NNPDF weights:

  pronounced downward shift – not compatible with MSTWCMS



Non-linear extension: MSTW2008

- Compare the cross sections after the reweighting

- The 68% error sets

  Hessian reweighting:

  A small downward shift, compatible with MSTWCMS

  Bayesian reweighting with NNPDF weights:

  pronounced downward shift – not compatible with MSTWCMS

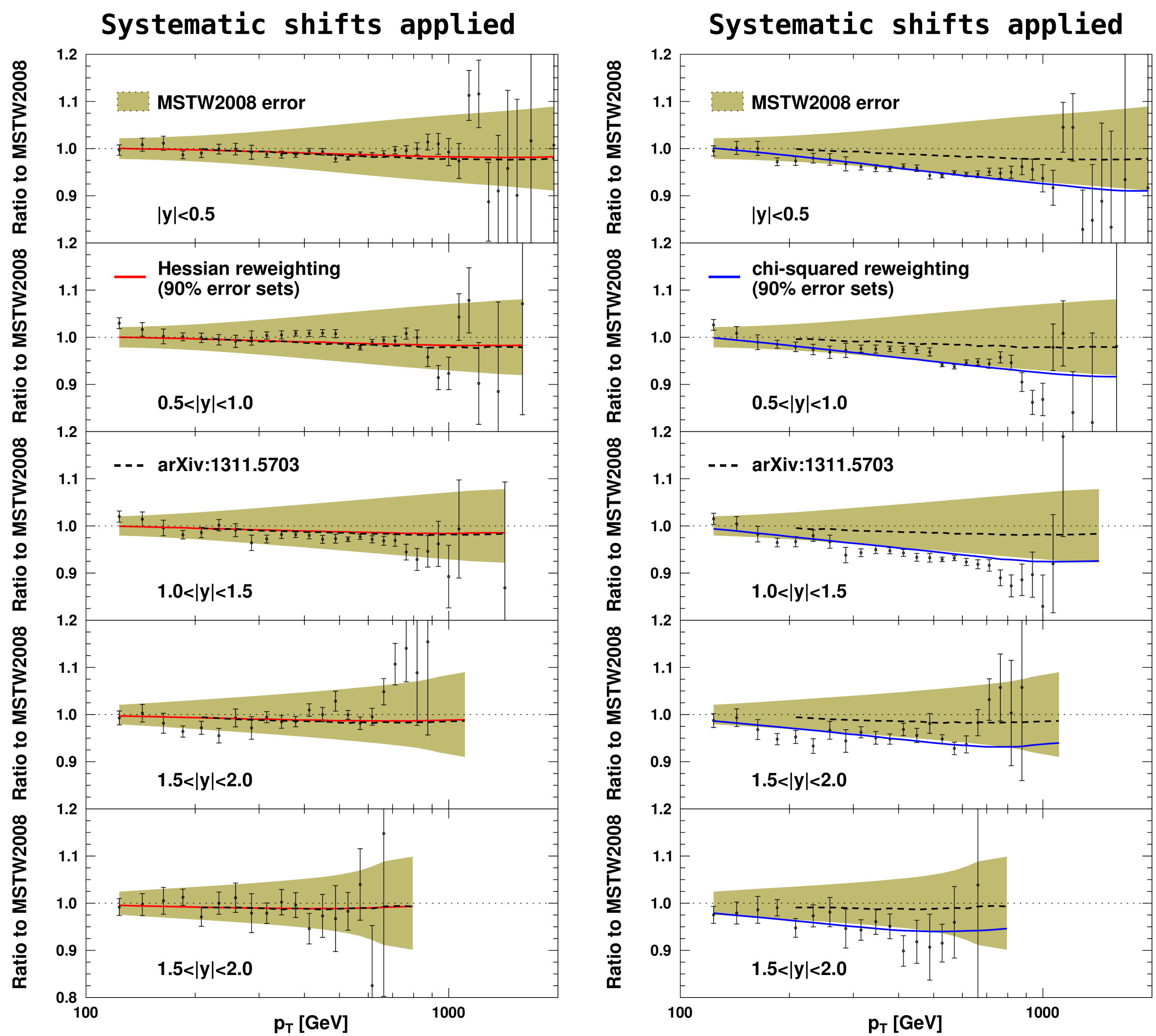
- The 90% error sets

  Hessian reweighting:

  As with 68% error sets

  Bayesian reweighting with NNPDF weights:

  Even more pronounced downward shift – not compatible with MSTWCMS



Summary

Baseline PDF fit

Hessian error sets
with non-zero tolerance

Hessian reweighting (linear/non-linear)

Direct re-fit

Rescaled Giele-Keller weights

$\chi^2 \longrightarrow \chi^2/\Delta\chi^2$

New set of PDFs

Tries to minimize the "global" χ^2

The result does not depend on $\Delta\chi^2$ (only affects the PDF errors)

NNPDF reweighting

New set of PDFs

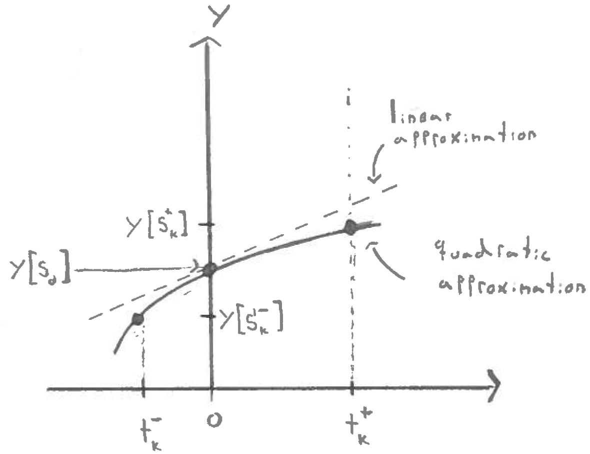
Tries to achieve $\chi^2/N=1$ for the new data

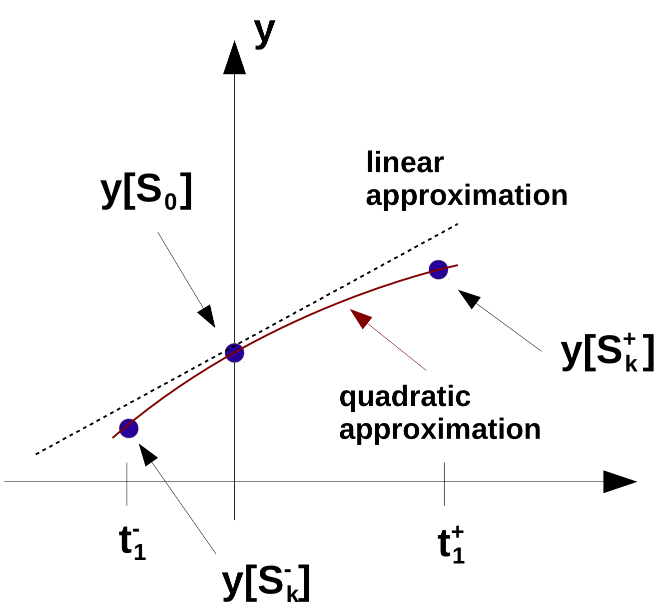
Assumes that the original replicas come from a native NNPDF fit

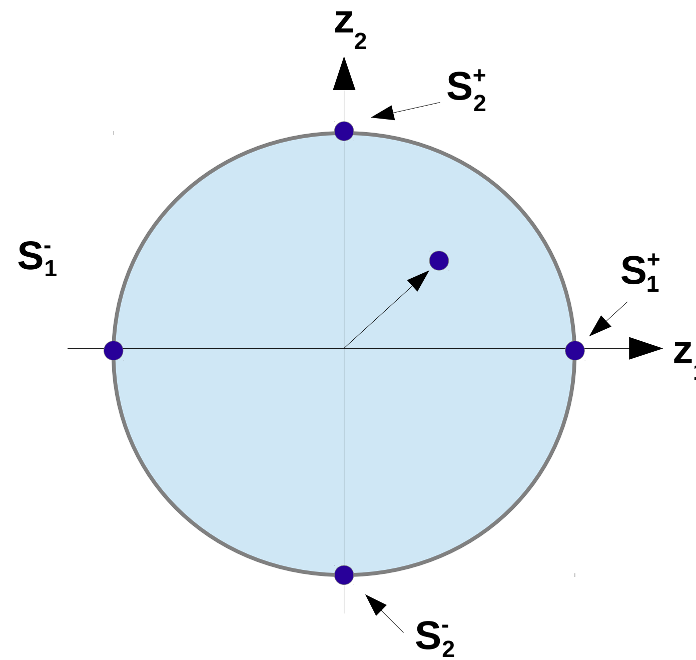
Result depend on $\Delta\chi^2$

NNPDF PDF fit : NNPDF reweighting

Hessian PDF fit : Hessian reweighting, rescaled Giele-Keller reweighting


Y
linear approximation
y[S_k^+]
y[S_0]
y[S_k^-]
quadratic approximation
t_k^-
0
t_k^+


y
linear approximation
y[S_0]
y[S_k^+]
quadratic approximation
t_1^-
t_1^+
y[S_k^-]


z_2
S_2^+
S_1^-
S_1^+
z_1
S_2^-