Works in the past week (2025/2/27)

- Cross check of CEM calculations of Jpsi production with 蔡燕兵。
- CEM/NRQCD calculations of difference of sigma(pi+)-sigma(pi-).
- Read papers about unbinned unfolding (PRL 133, 261803 (2024))
- Chi2 expression

References

1. Uncertainties of Predictions from Parton Distribution Functions I: the Lagrange Multiplier Method https://arxiv.org/abs/hep-ph/0101051
2. New Generation of Parton Distributions with Uncertainties from Global QCD Analysis https://arxiv.org/abs/hep-ph/0201195
3. Parton Distribution Benchmarking with LHC Data https://arxiv.org/abs/1211.5142
4. Fitting Parton Distribution Data with Multiplicative Normalization Uncertainties https://arxiv.org/abs/0912.2276
5. Extraction of unpolarized quark transverse momentum dependent parton distributions from Drell-Yan/Z-boson production https://arxiv.org/abs/1902.08474

Ref. 1

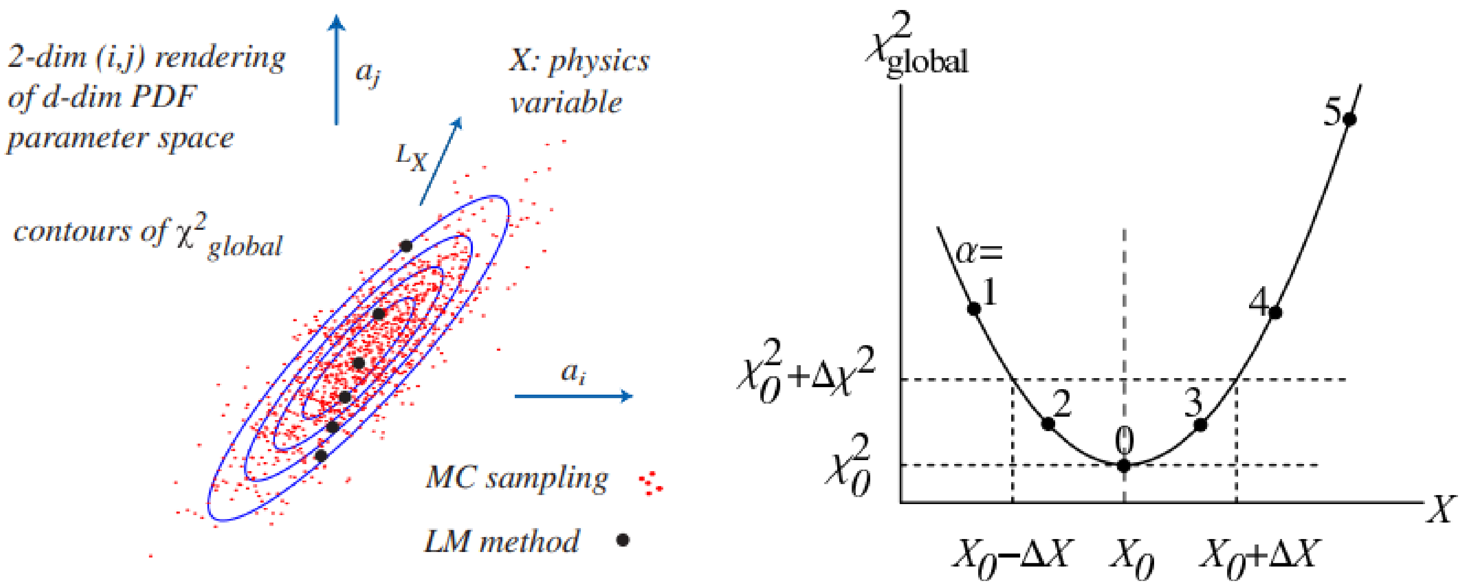


Figure 1: Left: The LM method provides sample points along a single curve L_X in the multi-dimensional PDF parameter space, relevant ro the observable X. Right: For a given tolerance $\Delta\chi^2_{\rm global}$, the uncertainty in the calculated value of X is $\pm\Delta X$. The solid points correspond to the sample points on the curve L_X in the left plot.

Ref. 2
Simplest case: statistical and uncorrelated systematic errors

The simplest χ^2 function, used in most conventional PDF analyses, is

$$\chi_0^2 = \sum_{\text{expt.}} \sum_{i=1}^{N_e} \frac{(D_i - T_i)^2}{\sigma_i'^2} \tag{6}$$

where D_i is a data value, T_i is the corresponding theory value (which depends on the PDF model parameters $\{a\}$), and σ_i' is the combined statistical and systematic errors (assumed uncorrelated and usually added in quadrature) on the measurement D_i. This effective χ^2 function provides a simple measure of goodness-of-fit, convenient for the search for candidate PDF sets by minimization. However, it is not useful for estimating the uncertainties associated with those candidates because it does not contain enough information to allow a meaningful statistical inference based on the increase in χ^2 away from the minimum.

Ref. 2 Real case: with correlated systematic error β_{ki}

Expression of chi2 with Penalty Terms

Most DIS experiments now provide more detailed information on measurement errors. For each data point i, we have the statistical error σ_i, uncorrelated systematic error u_i, and several (say, K) sources of correlated systematic errors $\{\beta_{1i}, \beta_{2i}, \ldots, \beta_{Ki}\}$. The best fit to the data (i.e., the fit with least variance) comes from minimizing the χ^2 function,[19]

$$\boxed{\chi'^2(\{a\},\{r\}) = \sum_{\text{expt.}} \left[\sum_{i=1}^{N_e} \frac{1}{\alpha_i^2} \left(D_i - T_i - \sum_{k=1}^{K} r_k \beta_{ki} \right)^2 + \sum_{k=1}^{K} r_k^2 \right]} \tag{7}$$

where $\boxed{\alpha_i^2 = \sigma_i^2 + u_i^2}$ is the combined uncorrelated error. The fitting parameters are (i) the PDF model parameters $\{a\}$, on which T_i depends, together with (ii) random parameters $\{r\}$ associated with the sources of correlated systematic error. The point of Eq. (7) is that D_i has a fluctuation $\sum_k r_k \beta_{ki}$ due to systematics. The best estimate of this shift is obtained by minimizing χ'^2 with respect to the set $\{r_k\}$. In practice, the total number of such parameters for all experiments

Ref. 2
Analytical Solution of minimizing chi2 w.r.t. {r}

We pointed out in Ref. [12] that the minimization of the function χ^2 with respect to $\{r\}$ can be carried out analytically. This simplifies the global analysis to its irreducible task of minimization with respect to the PDF parameters $\{a\}$ only. In addition, the analytic method provides explicit formulas for the optimal values of $\{r_k, k = 1 \ldots K\}$ due to the systematic errors $k = 1 \ldots K$ that are associated with the fit. These optimal shifts are obtained from the condition $\partial\chi^2/\partial r_k = 0$, and the result is

Analytical solution of {r}

$$\boxed{r_k(\{a\}) = \sum_{k'=1}^{K} \left(A^{-1}\right)_{kk'} B_{k'}} \,. \tag{8}$$

Here $\{B_k\}$ and $\{A_{kk'}\}$ are given by

$$B_k(\{a\}) = \sum_{i=1}^{N_e} \frac{\beta_{ki}\,(D_i - T_i)}{\alpha_i^2} \quad \text{and} \quad \boxed{A_{kk'} = \delta_{kk'} + \sum_{i=1}^{N_e} \frac{\beta_{ki}\beta_{k'i}}{\alpha_i^2}} \,. \tag{9}$$

Substituting the best estimates (8) back into χ'^2, we obtain a simplified χ^2 function,

$$\boxed{\chi^2(\{a\}) \equiv \chi'^2(\{a\}, \{r(\{a\})\}) = \sum_{\text{expt.}} \left\{ \sum_{i=1}^{N_e} \frac{(D_i - T_i)^2}{\alpha_i^2} - \sum_{k,k'=1}^{K} B_k \left(A^{-1}\right)_{kk'} B_{k'} \right\}} \,. \tag{10}$$

Ref. 2 Real case: with correlated systematic error β_{ki}

Expression of chi2 with Penalty Terms

$$\boxed{\chi'^2(\{a\},\{r\}) = \sum_{\text{expt.}} \left[\sum_{i=1}^{N_e} \frac{1}{\alpha_i^2} \left(D_i - T_i - \sum_{k=1}^{K} r_k \beta_{ki} \right)^2 + \sum_{k=1}^{K} r_k^2 \right]} \tag{7}$$

Eqs. (7) and (10) as follows:

$$\boxed{\hat{\chi}^2 \equiv \chi^2(\{\hat{a}\}) = \chi'^2(\{\hat{a}\},\{\hat{r}\}) = \sum_{\text{expt.}} \left[\sum_{i=1}^{N_e} \frac{\left(\widehat{D}_i - T_i\right)^2}{\alpha_i^2} + \sum_{k=1}^{K} \hat{r}_k^2 \right]}\ , \tag{11}$$

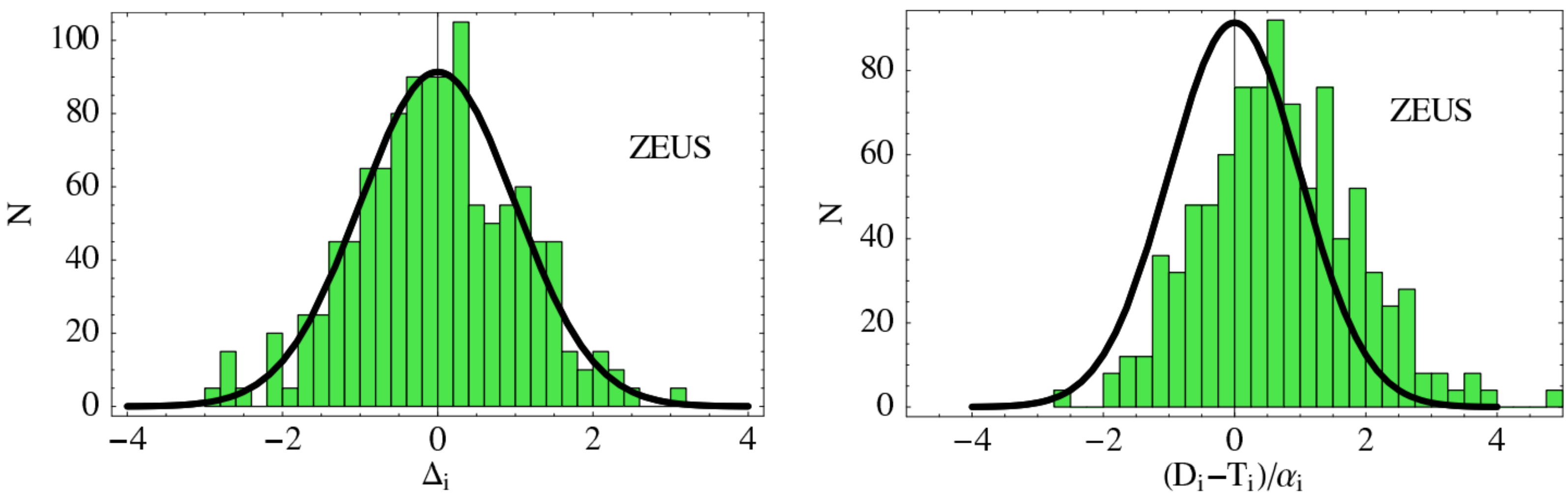
where $\{\hat{a}\}$ (like $\{\hat{r}\}$ above), are the PDF parameters $\{a\}$ at the χ^2 minimum, and

$$\widehat{D}_i \equiv D_i - \sum_{k=1}^{K} \hat{r}_k \beta_{ki}, \tag{12}$$

$$\hat{r}_k \ (\{a\}) = \sum_{k'=1}^{K} \left(A^{-1}\right)_{kk'} B_{k'} \ . \tag{8}$$

Ref. 2

Fig. 18 : (a) Histogram of Δ_i in Eq. (13) for the ZEUS data. The curve is a Gaussian of width 1. (b) A similar comparison but without the corrections for systematic errors on the data points.



$$\widehat{D}_i \equiv D_i - \sum_{k=1}^{K} \hat{r}_k \beta_{ki} \qquad \Delta_i \equiv \frac{\widehat{D}_i - T_i}{\alpha_i}$$

Ref. 3
Expression of chi2 with the covariance matrix

A.1 Definitions of χ^2 with the covariance matrix

We can define the χ^2 for a specific experiment with N_{pt} data points by

$$\boxed{\chi^2 = \sum_{i,j}^{N_{\mathrm{pt}}} (T_i - D_i)(\mathrm{cov}^{-1})_{ij}(T_j - D_j),} \tag{7}$$

and use it as a figure of merit to judge the agreement between theory and data. The covariance matrix $(\mathrm{cov})_{ij}$ used in this definition may be written as

$$\boxed{(\mathrm{cov})_{ij} = \delta_{ij} s_i^2 + \left(\sum_{\alpha=1}^{N_c} \sigma_{i,\alpha}^{(c)} \sigma_{j,\alpha}^{(c)} + \sum_{\alpha=1}^{N_{\mathcal{L}}} \sigma_{i,\alpha}^{(\mathcal{L})} \sigma_{j,\alpha}^{(\mathcal{L})} \right) D_i D_j,} \tag{8}$$

where i and j run over the experimental points $(i, j = 1, ..., N_{\mathrm{pt}})$, D_i are the measured central values, and T_i the corresponding theoretical predictions computed with a given set of PDFs. This covariance matrix depends on uncorrelated uncertainties s_i, constructed by adding the statistical and uncorrelated systematic uncertainties in quadrature; $N_{\mathcal{L}}$ multiplicative normalization uncertainties, $\sigma_{i,\alpha}^{(\mathcal{L})}$; and N_c other correlated systematic uncertainties, expressed for convenience in the above equation in terms of their relative values $\sigma_{i,\alpha}^{(c)}$. The total number of correlated uncertainties is thus $N_\lambda = N_{\mathcal{L}} + N_c$. Asymmetric systematic uncertainties provided by the experiments must be symmetrized to use this expression. We symmetrize them by averaging, $\sigma_{i,\alpha}^{(c)} = \frac{1}{2}(\sigma_{i,\alpha}^{(c),+} + \sigma_{i,\alpha}^{(c),-})$.

Ref. 3 Expression of chi2 with Penalty Terms

A.2 Definitions of χ^2 with shift parameters

An alternative, yet numerically equivalent, representation for the χ^2 function has been used in the jet benchmarking exercise of Sec. 5, following the method traditionally adopted in the CTEQ and MSTW PDF fits for jet and some other data sets. In this representation, the χ^2 figure of merit for goodness-of-fit to an experiment with correlated systematic uncertainties is expressed as [77]

$$\boxed{\chi^2(\{a\},\{\lambda\}) = \chi^2_D + \chi^2_\lambda,} \tag{11}$$

where

$$\chi^2_D \equiv \sum_{k=1}^{N_{\rm pt}} \frac{1}{s_k^2}\left(D_k - T_k - \sum_{\alpha=1}^{N_\lambda} \beta_{k,\alpha}\lambda_\alpha\right)^2, \tag{12}$$

Uncorrelated contributions

and

$$\chi^2_\lambda \equiv \sum_{\alpha=1}^{N_\lambda} \lambda_\alpha^2, \tag{13}$$

Penalty terms

using the same notation as in the previous section, where the $\beta_{k,\alpha}$ are the absolute correlated uncertainties. Systematic uncertainties associated with N_λ sources may now induce correlated variations (shifts) in the experimental data points. Their effect is approximated by including a sum $\sum_\alpha \beta_{k,\alpha}\lambda_\alpha$ dependent on the correlation matrix $\beta_{k,\alpha}$ ($k = 1, ..., N_{\rm pt}$; $\alpha = 1, ..., N_\lambda$) and stochastic nuisance parameters λ_α, with one nuisance parameter assigned to every source of the systematic uncertainty. By a common assumption, each λ_α follows the standard normal distribution. Its deviation from $\lambda_\alpha = 0$ incurs a penalty contribution λ_α^2 to χ^2. Under this assumption the minimum of χ^2 with respect to λ_α can be found algebraically, since the dependence on λ_α is quadratic [77].

Ref. 3 Matching of two chi2 expressions

We can solve for the best-fit values $\lambda_{0\alpha}$ of the nuisance parameters to find

$$\lambda_{0\alpha} = \sum_{i=1}^{N_{\rm pt}} \frac{D_i - T_i}{s_i} \sum_{\delta=1}^{N_\lambda} \mathcal{A}^{-1}_{\alpha\delta} \frac{\beta_{i,\delta}}{s_i}, \quad (14)$$

with

$$\boxed{\mathcal{A}_{\alpha\beta} = \delta_{\alpha\beta} + \sum_{k=1}^{N_{\rm pt}} \frac{\beta_{k,\alpha}\beta_{k,\beta}}{s_k^2}.} \quad (15)$$

When these $\lambda_{0\alpha}$ values are substituted into Eq. (13), one obtains the usual expression Eq. (7) for the χ^2, with

$$\boxed{({\rm cov})^{-1}_{ij} = \left[\frac{\delta_{ij}}{s_i^2} - \sum_{\alpha,\beta=1}^{N_\lambda} \frac{\beta_{i,\alpha}}{s_i^2} \mathcal{A}^{-1}_{\alpha\beta} \frac{\beta_{j,\beta}}{s_j^2}\right],} \quad (16)$$

the inverse of

$$\boxed{({\rm cov})_{ij} \equiv s_i^2 \delta_{ij} + \sum_{\alpha=1}^{N_\lambda} \beta_{i,\alpha}\beta_{j,\alpha}.} \quad (17)$$

$$\boxed{({\rm cov})_{ij} = \delta_{ij} s_i^2 + \left(\sum_{\alpha=1}^{N_c} \sigma^{(c)}_{i,\alpha}\sigma^{(c)}_{j,\alpha} + \sum_{\alpha=1}^{N_{\mathcal{L}}} \sigma^{(\mathcal{L})}_{i,\alpha}\sigma^{(\mathcal{L})}_{j,\alpha}\right) D_i D_j,} \quad (8)$$

$$\beta_{i,\alpha} = \sigma_{i,\alpha} D_i, \quad (18)$$