Systematic Errors in PDF/α_S Fits

Tim Bolton
Kansas State University

ABSTRACT

Procedures are described for reporting systematic errors in structure function determinations, and for using the reported information in PDF determinations from the data. Adoption of these procedures, or ones similar, by CTEQ and other global PDF fitters would signifigantly improve the utility of PDF parametrizations.

I. INTRODUCTION

Assume that an experiment N measures a structure function i in a bin centered at $\vec{x}_j$, $F_i^N(\vec{x}_j)$ with a statistical uncertainty $\Delta F_i^N(\vec{x}_j)$. One would like to fit a model to these measured structure functions for this and other experiments, and extract physics, typically a value of the strong coupling constant at $Q^2 = M_Z^2$, or equivalently, $\Lambda_{\overline{MS}}$, and a set of numbers, $\vec{P}$, parameterizing parton distributions at a reference scale Q_0^2. The set of numbers $\vec{P}$ include parton distribution function (PDF) parameters, as well as auxiliary quantities needed by QCD, such as the renormilization scale μ or a heavy quark mass m_Q.

The model structure function for experiment N will also depend on a variety of corrections for calibrations, acceptance, normalization, and other factors; uncertainties in these corrections lead to systematic errors. This paper describes a method of incorporating effects of systematic errors into the determination of $\vec{P}$ and $\Lambda_{\overline{MS}}$.

II. INCORPORATING SYSTEMATIC EFFECTS DIRECTLY INTO MODEL

The set of systematic effects $\epsilon_l^N(\vec{x}_j)$ can be regarded as a part of the model. One can describe the systematic uncertainties by the definition

$$\epsilon_l^N(\vec{x}_j) = \bar{\epsilon}_l^N(\vec{x}_j) + \Delta\epsilon_l^N(\vec{x}_j), \tag{1}$$

where the nominal value $\bar{\epsilon}_l^N(\vec{x}_j)$ can be arranged to be zero and the unknown deviation from the nominal value $\Delta\epsilon_l^N(\vec{x}_j)$ is constrained by other information (test beams, monte carlos, theoretical prejudice, etc.) to

$$\frac{\Delta\epsilon_l^N(\vec{x}_j)}{\delta\epsilon_l^N(\vec{x}_j)} \leq 1, \tag{2}$$

with $\delta\epsilon_l^N(\vec{x}_j)$ the estimated uncertainty in the particular systematic effect (e.g., a 1% normalization error, a 2.5% energy scale error, etc.)

A given experiment can measure or calculate $\delta\epsilon_l^N(\vec{x}_j)$ for each bin j and each systematic error source l, and, in principle, all the correlations between the different i, l combinations. Generally this is unnecessary because:

- One can generally choose a set of uncorrelated systematic errors by combining sets of highly correlated quantities.
- The kinematic dependence of systematic effects often factors into the form

$$\Delta\epsilon_l^N(\vec{x}_j) = \hat{\epsilon}_l^N f_l^N(\vec{x}_j), \tag{3}$$

  with $f_l^N(\vec{x}_j)$ a known shape function and $\hat{\epsilon}_l^N$a constant. With this form, the correlations across kinematic variables are incorporated automatically. Experiments can provide the uncertainty in $\hat{\epsilon}_l^N$, $\delta\hat{\epsilon}_l^N$, and the shape function $f_l^N(\vec{x}_j)$; and by appropriate choice of the scale of $f_l^N(\vec{x}_j)$, it is possible to arrange for $\delta\hat{\epsilon}_l^N = 1$.

III. LINEARIZED MODEL

A form of the structure function suitable for fitting can be easily obtained by expanding about a nominal model

$$\begin{aligned}\tilde{F}_i^N(\vec{x}_j;\Lambda_{\overline{MS}},\vec{P};\vec{\epsilon}^N(\vec{x}_j)) &= \tilde{F}_i^N(\vec{x}_j;\Lambda_{\overline{MS}0},\vec{P}_0;0) \\ &+\frac{\partial\tilde{F}_i^N(\vec{x}_j;\Lambda_{\overline{MS}0},\vec{P}_0;0)}{\partial\Lambda_{\overline{MS}}}\Delta\Lambda_{\overline{MS}} \\ &+\frac{\partial\tilde{F}_i^N(\vec{x}_j;\Lambda_{\overline{MS}0},\vec{P}_0;0)}{\partial\vec{P}}\Delta\vec{P} \\ &+\sum_l\frac{\partial\tilde{F}_i^N(\vec{x}_j;\Lambda_{\overline{MS}0},\vec{P}_0;0)}{\partial\epsilon_l^N(\vec{x}_j)} \\ &\times f_l^N(\vec{x}_j)\hat{\epsilon}_l^N.\end{aligned} \tag{4}$$

This form is then used directly in the χ^2 function

$$\chi_N^2 = \sum_j\left(\frac{F_i^N(\vec{x}_j) - \tilde{F}_i^N(\vec{x}_j;\Lambda_{\overline{MS}},\vec{P};\vec{\epsilon}^N(\vec{x}_j))}{\Delta F_i^N(\vec{x}_j)}\right)^2 + \sum_l\left(\hat{\epsilon}_l^N\right)^2 \tag{5}$$

In this expression:

- $\tilde{F}_i^N(\vec{x}_j;\Lambda_{\overline{MS}0},\vec{P}_0;0)$ is a nominal, preferably "close", model of the structure functions. It is defined by initial best guess choices for the QCD scale, $\Lambda_{\overline{MS}0}$, and PDF parameters $\vec{P}_0$; and assumes zero systematic effects from all other theoretical or experimental sources. $\tilde{F}_i^N(\vec{x}_j;\Lambda_{\overline{MS}0},\vec{P}_0;0)$ is defined by whoever wants to fit the structure functions and requires no experimental input.
- $\frac{\partial\tilde{F}_i^N(\vec{x}_j;\Lambda_{\overline{MS}0},\vec{P}_0;0)}{\partial\Lambda_{\overline{MS}}}$ and $\frac{\partial\tilde{F}_i^N(\vec{x}_j;\Lambda_{\overline{MS}0},\vec{P}_0;0)}{\partial\vec{P}}$ define dependencies of the model structure functions on $\Lambda_{\overline{MS}}$ and PDF parameters $\vec{P}$. These numbers are evaluated for the nominal model and require no input from experiment.

- $\frac{\partial \tilde{F}_i^N(\vec{x}_j;\Lambda_{\overline{MS}0},\vec{P}_0;0)}{\partial \epsilon_l^N(\vec{x}_j)}$ and $f_l^N(\vec{x}_j)$ are the sets of numbers that define the sensitivity of the model to experimental and systematic uncertainties. In general, the experiment must provide these; and there are at least two methods that could be used to convey the information:

  - $\frac{\partial \tilde{F}_i^N(\vec{x}_j;\Lambda_{\overline{MS}0},\vec{P}_0;0)}{\partial \epsilon_l^N(\vec{x}_j)}$ $f_l^N(\vec{x}_j)$ could be provided in a table for each systematic effect l at each measured kinematic point $\vec{x}_j$.
  - Alternatively, the structure function evaluated for a specified (usually 1σ) shift in the parameter $\hat{\epsilon}_l^N$ corresponding to systematic effect l at each measured kinematic point $\vec{x}_j$ could be given. This second form would in general convey more information; for example, structure function values for $\pm\hat{\epsilon}_l^N$ allow a check for asymmetric errors[1].

- The part of the χ^2 function $\sum_l \left(\hat{\epsilon}_l^N\right)^2$, a "penalty function", allows the fit to incorporate the extra information embodied in the known $\delta\hat{\epsilon}_l^N$.

- $\Delta\Lambda_{\overline{MS}}$, $\Delta\vec{P}$, and the $\hat{\epsilon}_l^N$ are fit parameters. It is the inclusion of the latter set as free parameters that makes this approach to fitting different, although actually one is only extending a common practice of letting one systematic error parameter, the normalization, float.

Minimization of the fit through direct matrix inversion or by use of MINUIT [2] then produces the following information:

- The usual best estimates for $\Delta\Lambda_{\overline{MS}}$ and $\Delta\vec{P}$, which yield the desired physical parameters via $\Lambda_{\overline{MS}0} + \Delta\Lambda_{\overline{MS}}$ and $\vec{P}_0 + \Delta\vec{P}$.

- Estimates of the systematic error parameters $\hat{\epsilon}_l^N$ chosen so that the model best describes the data and is consistent with the known $\delta\hat{\epsilon}_l^N$

- A covariance matrix which can be inverted to give the full error matrix among all fit parameters.

- A reasonable "χ^2" value. Since $\sum_l \left(\hat{\epsilon}_l^N\right)^2$ will not be a true χ^2 term, the definition is not perfect. However, several advantages exist over the usual technique of adding systematic errors in quadrature with statistical errors:

  - Data are weighted by an understood error term. Hence, data from different experiments will combine in a more correct fashion.
  - Correlations between "physics" parameters and "systematic" parameters are automatically taken into account by the procedure.
  - A test of the "reasonableness" of systematic error estimates follows from observing the pulls on the parameters $\hat{\epsilon}_l^N$.

IV. COMMENTS

A. What should come out?

The previous section defines precisely the way experiments should provide information about systematic errors and the way PDF fitters should use this information. The fitters should provide, preferably as computer code:

- A complete description of the PDF model and its assumptions. (Done now).

- The PDF's as a function of x and Q^2, $\vec{q}(x,Q^2)$, and $\alpha_S(Q^2)$, in a well-defined scheme. (Done now).

- A set of covariance matrices. (New) To be precise, one wants computer code that returns (the sum over repeated index is implied):

$$\begin{aligned}
&\left\langle \delta q_i(x_1,Q_1^2;\Lambda_{\overline{MS}},\vec{P})\delta q_j(x_2,Q_2^2;\Lambda_{\overline{MS}},\vec{P})\right\rangle \\
=\ & \frac{\partial q_i(x_1,Q_1^2)}{\partial P_m}\frac{\partial q_j(x_2,Q_2^2)}{\partial P_n}\left\langle \delta P_m \delta P_n\right\rangle \\
&+\left(\frac{\partial q_i(x_1,Q_1^2)}{\partial P_m}\frac{\partial q_j(x_2,Q_2^2)}{\partial \Lambda_{\overline{MS}}}+\frac{\partial q_j(x_1,Q_1^2)}{\partial P_m}\frac{\partial q_i(x_2,Q_2^2)}{\partial \Lambda_{\overline{MS}}}\right) \\
&\times\left\langle \delta P_m \delta\Lambda_{\overline{MS}}\right\rangle \\
&+\frac{\partial q_i(x_2,Q_2^2)}{\partial \Lambda_{\overline{MS}}}\frac{\partial q_j(x_2,Q_2^2)}{\partial \Lambda_{\overline{MS}}}\left\langle \delta\Lambda_{\overline{MS}}\delta\Lambda_{\overline{MS}}\right\rangle ;
\end{aligned} \tag{6}$$

and

$$\begin{aligned}
&\left\langle \delta\alpha_S(Q_1^2;\Lambda_{\overline{MS}})\delta q_i(x_2,Q_2^2;\Lambda_{\overline{MS}},\vec{P})\right\rangle \\
=\ & \frac{\partial\alpha_S(Q_1^2;\Lambda_{\overline{MS}})}{\partial\Lambda_{\overline{MS}}}\frac{\partial q_i(x_2,Q_2^2)}{\partial\Lambda_{\overline{MS}}}\left\langle \delta\Lambda_{\overline{MS}}\delta\Lambda_{\overline{MS}}\right\rangle \\
&+\frac{\partial\alpha_S(Q_1^2;\Lambda_{\overline{MS}})}{\partial\Lambda_{\overline{MS}}}\frac{\partial q_i(x_2,Q_2^2)}{\partial P_m}\left\langle \delta P_m\delta\Lambda_{\overline{MS}}\right\rangle ;
\end{aligned} \tag{7}$$

and

$$\begin{aligned}
&\left\langle \delta\alpha_S(Q_1^2;\Lambda_{\overline{MS}})\delta\alpha_S(Q_2^2;\Lambda_{\overline{MS}})\right\rangle \\
=\ & \frac{\partial\alpha_S(Q_1^2;\Lambda_{\overline{MS}})}{\partial\Lambda_{\overline{MS}}}\frac{\partial\alpha_S(Q_2^2;\Lambda_{\overline{MS}})}{\partial\Lambda_{\overline{MS}}}\left\langle \delta\Lambda_{\overline{MS}}\delta\Lambda_{\overline{MS}}\right\rangle ;
\end{aligned} \tag{8}$$

where the PDF fit returns the more fundamental error matrices $\langle\delta P_m\delta P_n\rangle$, $\langle\delta P_m\delta\Lambda_{\overline{MS}}\rangle$, and $\langle\delta\Lambda_{\overline{MS}}\delta\Lambda_{\overline{MS}}\rangle$. These covariance matrices should allow a more proper error estimate for any cross section constructed from the PDF and α_S. It should be straightforward to implement computer code that performs these calculations; the simplest scheme would be a tabulation on a finite grid of points with interpolation.

- The (x,Q^2) range, χ_N^2 , number of systematic parameters, and the distribution of $\hat{\epsilon}_l^N$ for each experiment used in the fit.

B. How to use incomplete information

Older experiments may not have produced a detailed tabulation of systematic errors. As long as the total systematic error at each point is presented, one can still attempt to follow the procedure outlined in this note. There can only be one systematic error parameter, $\hat{\epsilon}_1^N$, with $\delta\hat{\epsilon}_1^N = 1$, and $f_1^N(\vec{x})$ defined to be equal to the systematic error at each point $\vec{x}_j$; this amounts to assuming that the systematic errors are totally correlated from bin to bin. One then attempts to fit all the data by adjusting the physics model parameters and the single systematic error parameter.

Such a fit will likely have a poor χ^2 if the hypothesis of total correlation of all systematic errors is wrong since there are relatively fewer parameters to fit the data. Two possible choices in such an outcome are:

- Throw the experiment out.
- Keep the experiment, but "punish" its omission of systematic error information by inflating all uncorrelated error contributions by the factor $\sqrt{\frac{\chi_N^2}{N_{DOF}}}$, with N_{DOF} the number of degrees of freedom in the fit from the experiment, and then re-fitting.

Neither of these options is terribly satisfying; the safest course would be to include only experiment results with full systematic error tables into PDF fits.

C. QCD as a Calibration Tool

One, perhaps unsettling, side effect of the fitting procedure described here is that very high quality (low statistical error) data fitted with a QCD plus systematic error model can produce improvements in the systematic error parameters. As a crude example, consider the case of fitting M measurements, Y_m, of a y distribution to the theoretical form $A + B(1-y)^2$. Assume further that a calibration uncertainty introduces a systematic effect that can be described by $C(1-y)$, with C known to be less than 1%. The fit χ^2 would take the form

$$\chi^2 = \sum_{m=1}^{M} \left(\frac{Y_m - A - B(1-y_m)^2 - C(1-y_m)}{\Delta Y_m} \right)^2 + \frac{C^2}{0.01^2}. \tag{9}$$

If the statistics became very high, the parameter C would be determined better by fitting to the data than it would from the "external calibration", and systematic errors would be reduced by statistics. Matters would be different of course if the physics model contained a piece $D(1-y)$. In general, systematics can be improved by systematics if the systematic error contribution differs in "shape" from the physics model.

D. Number of Parameters

If each experiment N used by a PDF fit has L_N parameters, and there are $M+1$ fit parameters for $\vec{P}$ and the strong coupling constant, then an overall PDF fit must determine $M + 1 + \sum_{expts} L_N$ numbers. If the problem is linearized, as in this paper, then the problem reduces to inverting a large matrix; and many numerical analysis routines will do this. If MINUIT is used, then one might have to expand the size of some arrays.

V. REFERENCES

[1] For a recent example of this method, see *Measurement of the F_2 Structure Function in Deep Inelastic e^+p Scattering Using 1994 Data from The ZEUS Detector at HERA*, M. Derrick *et al* (ZEUS Collaboration), DESY preprint DESY 96-076 (1996).

[2] F. James, *MINUIT Reference Manual*, CERN Program Library Long Writeup D506 V94.1 (1994).