

Correlated uncertainties: xFitter

Oleksandr Zenaiev (DESY)
for xFitter developers team

12.07.2018

`xfitter.org`

What is xFitter?

- xFitter is open-source QCD fit framework
- xFitter is widely used for phenomenological studies to assess impact of new data and/or new theoretical developments

How it works (briefly)?

- specify which data and theory to use and what to do: fit, PDF profiling etc.
- data and theory are compared and χ^2 is calculated
 - *crucial to have proper data uncertainties*
 - sometimes feedback to the analysers is provided, if tensions are seen, and they may update the correlation model. Example: xFitter paper on Tevatron W,Z data [Eur.Phys.J. C75 (2015) 458]
- results are obtained, plotted, stored in LHAPDF format etc.

Treatment of correlated uncertainties in xFitter

Two equivalent modes are supported in xFitter:

- **Covariance matrix**
 - ▶ natural e.g. for unfolding uncertainties or normalised cross sections
 - ▶ possible to account for correlations across data sets (rarely provided/used)
- **Nuisance parameters**
 - ▶ natural for systematic uncertainties controlled by a single parameter
 - ▶ straightforward to be used to account for correlations across many data sets

Presently, **simple text format** is used to provide uncertainties in xFitter:

- most of data sets were implemented in xFitter by original authors when doing phenomenological analysis of their data
 - ▶ example: HERA data [Eur.Phys.J. C75 (2015) 580, ...]
 - ▶ some data tables are presently made public **only** in xFitter format: Eur.Phys.J. C78 (2018) 473 (recent HERA charm and beauty data)
- rarely we need to access something from papers or HepData
 - ▶ example: LHCb data on charm and beauty production [Nucl.Phys. B871 (2013) 1, JHEP 1308 (2013) 117]
 - *it would be beneficial to have a unified format*
 - ... but LHCb data were provided without correlations :-(
 - *first we need to convince experiments **always** to give information on correlations such that their data are used properly and most efficiently!*

Suggestions from xFitter concerning the format of correlated uncertainties:

- appreciate if two modes (covariance matrix and nuisance parameters) are supported
 - although equivalent, each of them has special features
- perhaps specific format is not really important
 - ▶ providing it will not be broken one day (or start to require a new library etc.)
 - ▶ personally, I prefer simple text format, but have no problem with other formats as well

Some questions to discuss:

- different treatment of uncertainties (such as multiplicative vs. additive) can be very important for precise data
- moreover, in some cases only certain treatment makes sense (e.g. for normalised cross sections uncertainties should be treated as additive, and one bin has to be dropped)
 - (?) should also this kind of information be provided with uncertainties?
 - (?) should it be part of the format? (but it may be good to keep a possibility for simplest treatment, i.e. additive, for some applications)