Investigations on unfolding

with toy distributions

Patrick L.S. Connor

Deutsches Elektronen-Synchrotron

12 October 2016

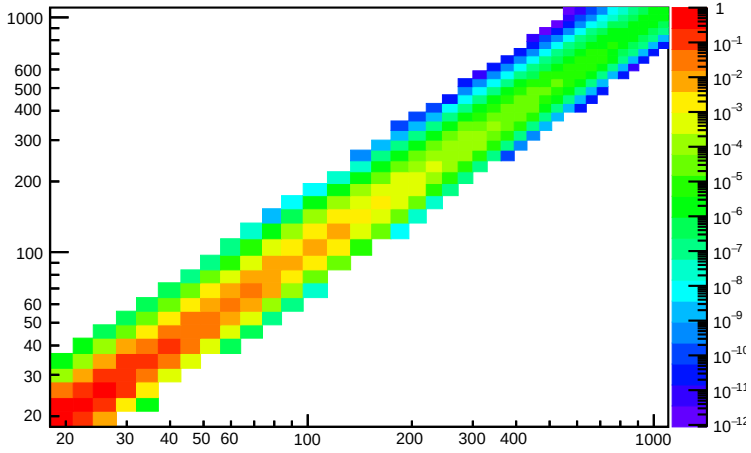
Tool

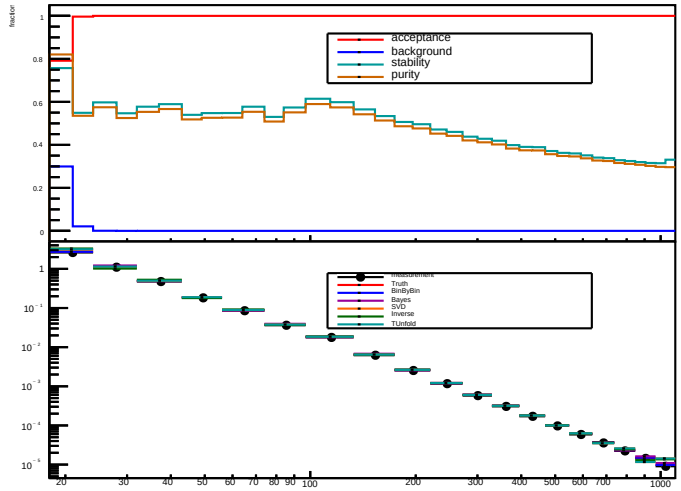
- standalone tool, compiled with the latest `RooUnfold`[1] and with `Root6`
  - code can be shared
- inputs are:
  - resolution (just a gaussian for the moment)
  - *true* p_T cross section for the measurement
  - *predicted* p_T cross section for the MC (+ construction of the RM)
  - binning scheme
- outputs are:
  - RM
  - differential resolution
  - ABPS
  - measured, true and unfolded spectra
  - ratios
- output format:
  - `.root` file
  - PDF files

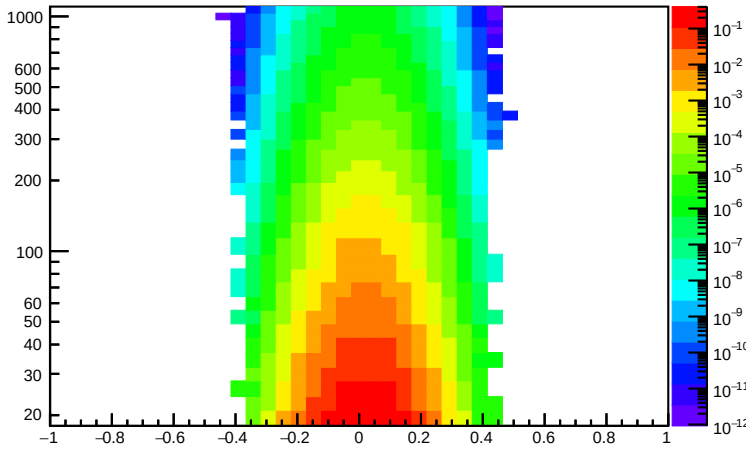
Questions

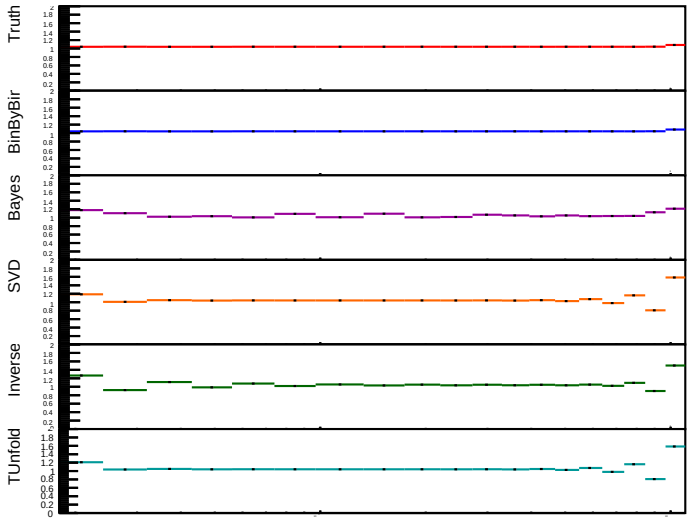
1. method (Bayes, inversion, ...)
2. p_T spectrum and model dependence
   $\longrightarrow$ looks like Panos' code is strongly model-dependent (?)
3. binning scheme + ABPS + miss/fake $\longrightarrow$ effect can be mostly seen on purity and stability
4. statistics + sampling (uniform/core)

Example









Many similar plots to investigate the impact of the different parameters (two series attached on the same indico slot)

Early conclusions I

- Different xsecs for the building of the RM and for the "truth"
  $\longrightarrow$ Bayes fails when different spectra are used (but could be due to fake/miss handling or to lack of statistics)
- Testing the two sampling methods (i.e. testing model dependence in the RM)
  $\longrightarrow$ Bayes works only with core sampling but all other work with both samplings
- Statistics
  $\longrightarrow$ not much difference between 10^7 and 10^8 except for the inversion method that can easily show "waves" (however, the sampling method and the xsec do matter)

$\longrightarrow$ looks that Bayes is very sensitive to the statistics and to the way the toy RM is filled

Early conclusions II

- Binning scheme
  $\longrightarrow$ standard binning works better than fixed-width binning (especially for Bayes)
- Checking ABPS
  $\longrightarrow$ fluctations appears when purity and stability are too low ($\tilde{\text{b}}$elow 1 σ)
  $\longrightarrow$ could explain the fluctuations when unfolding true data

$\longrightarrow$ larger bins could solve the fluctuations after unfolding on real data with Bayes

The End

Unfolding methods

bin/bin	model-dependent and wrong handling of statistical uncertainties
Bayes	iterative procedure, good experience in SMP, is shown to converge but unknown number of iterations
Inversion	best on principle, but possible instabilities if statistics is too low
TUnfold	likelihood minimisation, including regularisation, recommended by statistics comitee, developer is at DESY
SVD	...