

Unfolding: available methods, points to consider

CMS annual TOP PAG meeting 21 Nov 2019, DESY

Olaf Behnke (DESY)

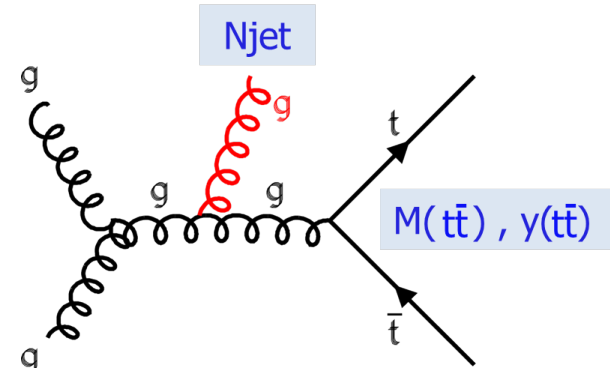
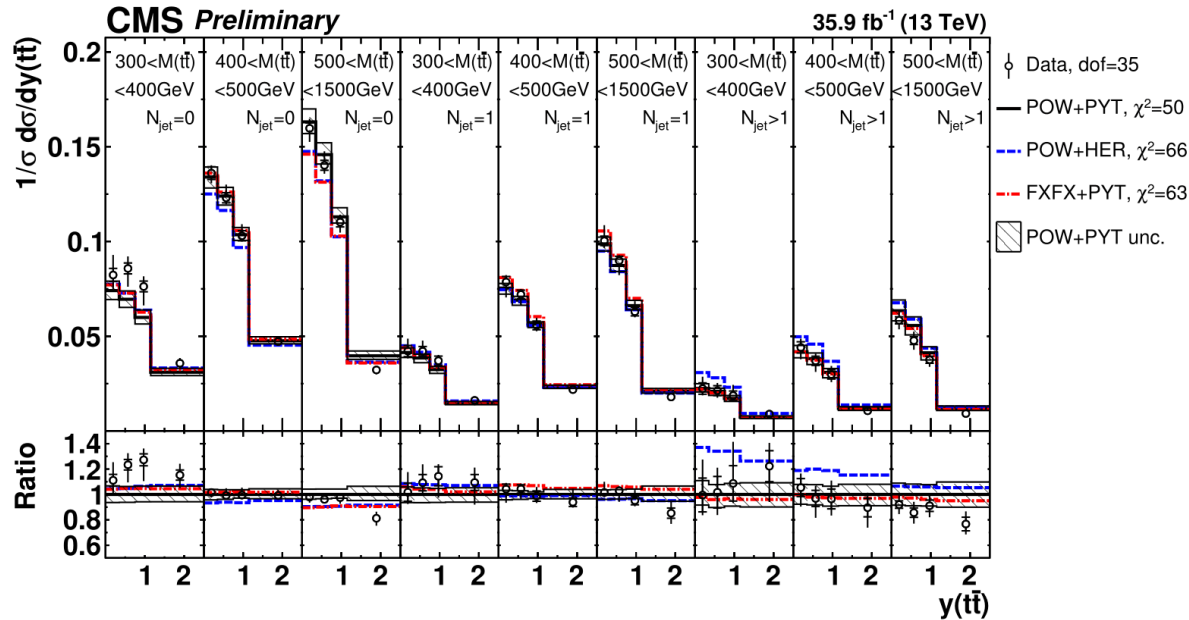
Tip: recent overview talk on unfolding methods in HEP by
M. Kuusela

<https://indico.desy.de/indico/event/22731/session/5/contribution/24/material/slides/0.pdf>

Introduction

- Measured already 3D $t\bar{t}$ xsecs with 2016 data:

TOP-18-004, 1904.05237



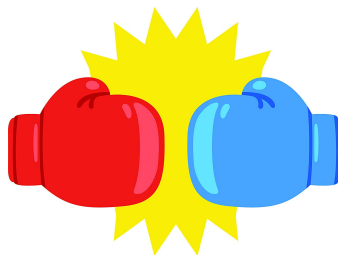
using TUnfold

Today:

Least squares **TUnfold**

<http://www.desy.de/~sschmitt/tunfold.html>

<http://iopscience.iop.org/1748-0221/7/10/T10003/>

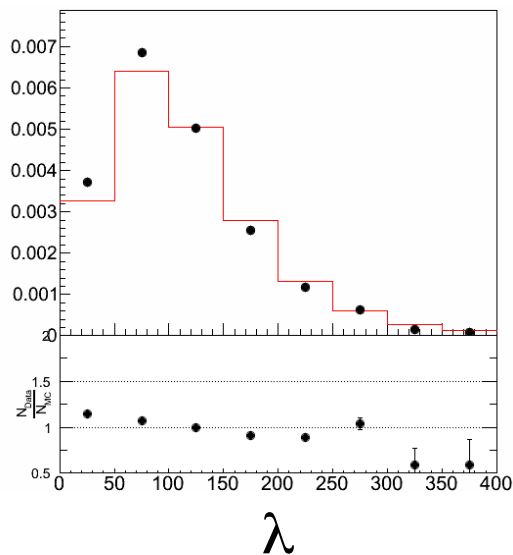


Max. Likelihood **Combine** tool

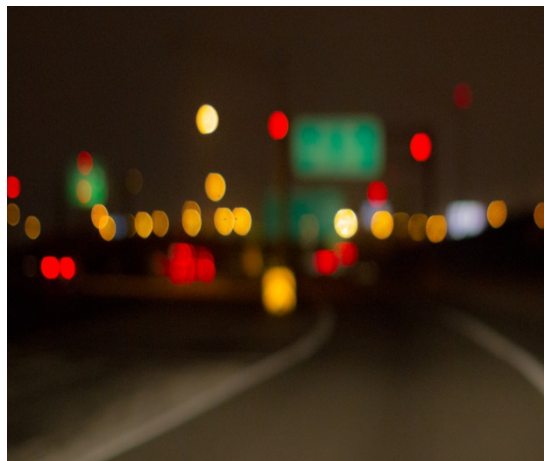
<https://cms-analysis.github.io/HiggsAnalysis-CombinedLimit>

Unfolding of λ

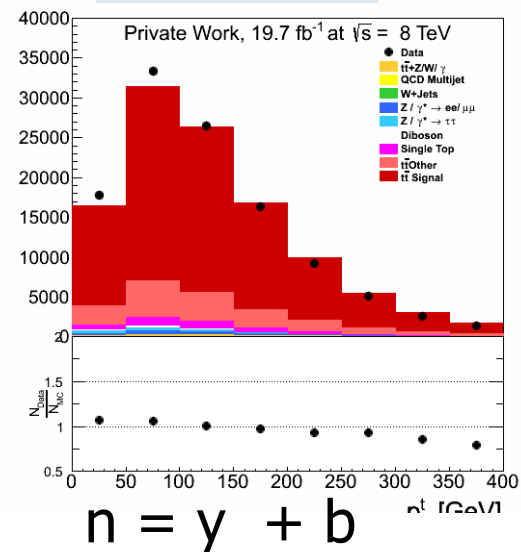
True level



Detector effects + background

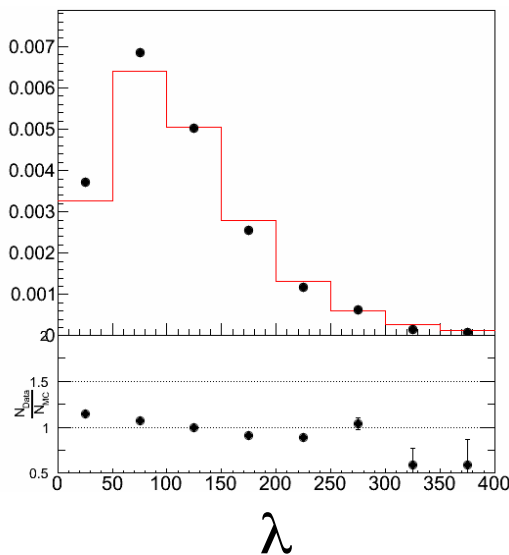


Smeared space

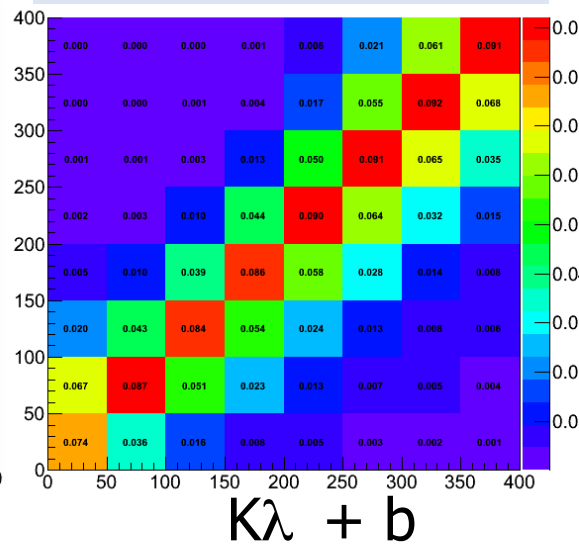


Unfolding of λ

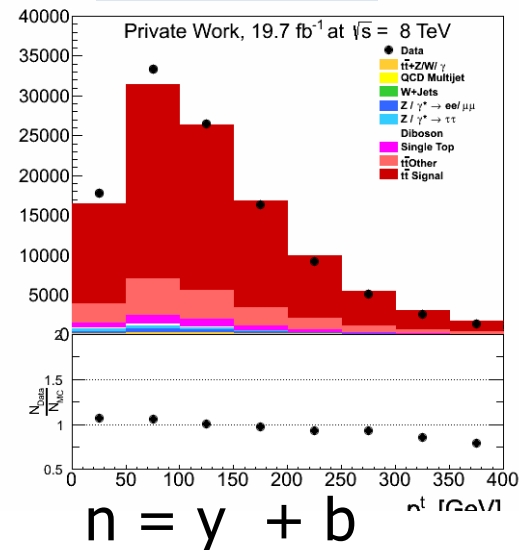
True level



Detector effects + background



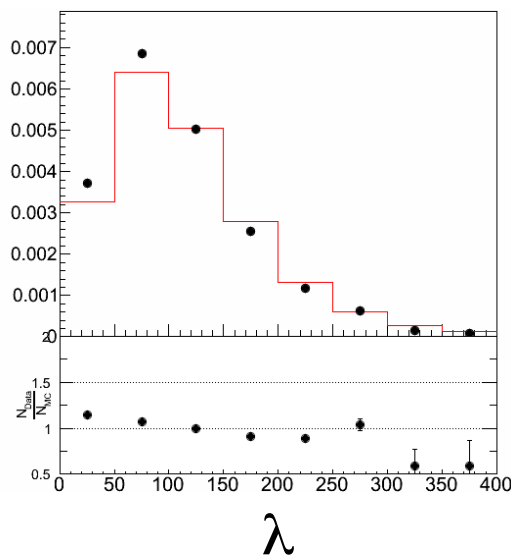
Smeared space



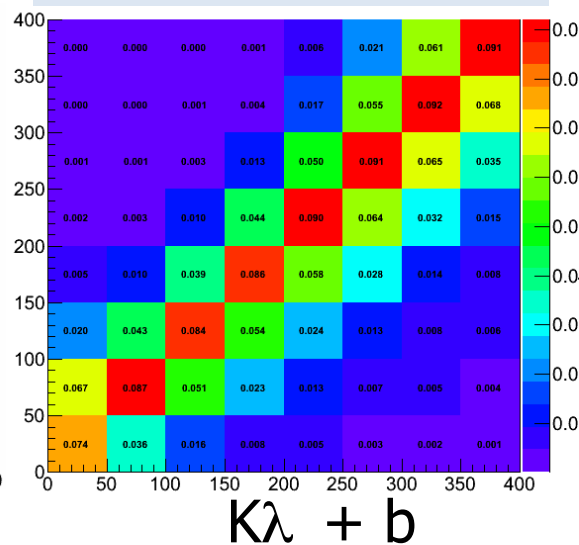
- Unfolding means estimating λ from n , $\rightarrow \hat{\lambda}$

Unfolding of λ

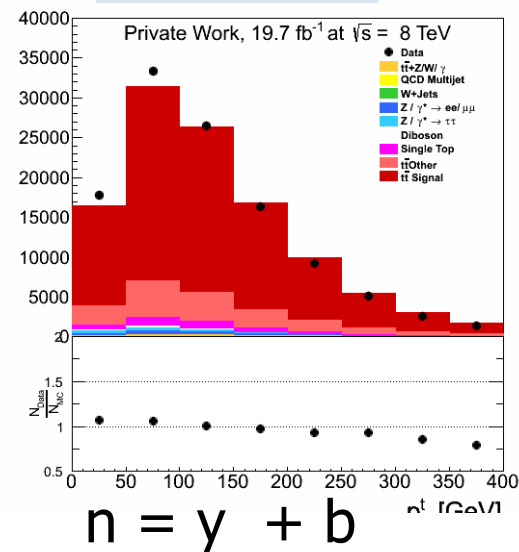
True level



Detector effects + background



Smeared space



Unfolding means estimating λ from n , $\rightarrow \hat{\lambda}$

TUnfold

$$\chi^2 = [\mathbf{y} - \mathbf{K}\boldsymbol{\lambda}]^t \mathbf{V}_y^{-1} [\mathbf{y} - \mathbf{K}\boldsymbol{\lambda}]$$

Combine

$$L \sim \prod_{j-\text{bins}} \text{Pois}[n_j | (\mathbf{K}\boldsymbol{\lambda})_j + b_j]$$

$$\tilde{\chi}^2 = -2 \ln L(\boldsymbol{\lambda})$$

Speed

TUnfold

$$\chi^2 = [\mathbf{y} - \mathbf{K}\boldsymbol{\lambda}]^t \mathbf{V}_y^{-1} [\mathbf{y} - \mathbf{K}\boldsymbol{\lambda}]$$

$$\text{Min. } \chi^2: d\chi^2/d\boldsymbol{\lambda}^t = 0$$

$$\Rightarrow \hat{\boldsymbol{\lambda}} = (\mathbf{K}^t \mathbf{V}_y^{-1} \mathbf{K})^{-1} \mathbf{V}_y^{-1} \mathbf{K}^t \mathbf{y}$$

Requires inversion of Hesse matrix:

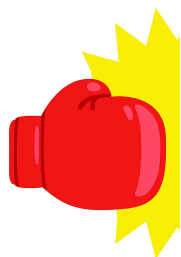
$$(\mathbf{K}^t \mathbf{V}_y^{-1} \mathbf{K})^{-1} = V_{\hat{\boldsymbol{\lambda}}}$$

Combine

$$L \sim \prod_{j-\text{bins}} \text{Pois}[n_j | (\mathbf{K}\boldsymbol{\lambda})_j + b_j]$$

$$\tilde{\chi}^2 = -2 \ln L(\boldsymbol{\lambda})$$

Numerical minimisation with Minuit → need thousands of calculations of $\tilde{\chi}^2$ for varying values of unfolded vector $\boldsymbol{\lambda}$



1 : 0

Precision

TUnfold

$$\chi^2 = [\mathbf{y} - \mathbf{K}\lambda]^t \mathbf{V}_y^{-1} [\mathbf{y} - \mathbf{K}\lambda]$$

Used Neyman χ^2 has known normalisation bias:

$$\#events_{\text{fitted}} = \#events_{\text{observed}} - \chi^2_{\text{min}}$$

TUnfold remedy: Area Constraint option

$$\chi^2 \rightarrow \chi^2 + \gamma \left(\sum_i (y_i - \sum_j K_{ij} \lambda_j) \right)$$

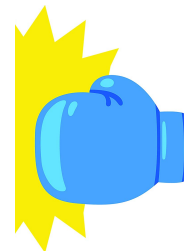
Lagrange multiplier

Combine

Poisson based likelihood yields

$$\#events_{\text{fitted}} = \#events_{\text{observed}}$$

1 : 1



Data Combination

TUnfold

Combine run II data $i=16, '17$ and $'18$

Uncorrelated yield vectors y^i :

$$\chi^2 = \sum_i [y_i - K_i \lambda]^t V_{y_i}^{-1} [y_i - K_i \lambda]$$

Correlated yield vectors y^i :

$$y = \begin{pmatrix} y^{16} \\ y^{17} \\ y^{18} \end{pmatrix} V_y = \begin{pmatrix} V^{16,16} & V^{16,17} & V^{16,18} \\ V^{16,17} & V^{17,17} & V^{17,18} \\ V^{16,18} & V^{17,18} & V^{18,18} \end{pmatrix} K = \begin{pmatrix} K^{16} \\ K^{17} \\ K^{18} \end{pmatrix}$$

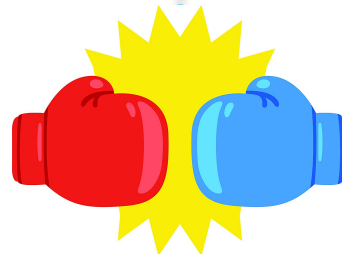
Min. $\chi^2 = [y - K\lambda]^t V_y^{-1} [y - K\lambda]$

Combine

Poisson counts n_j^i :

$$L \sim \prod_i \prod_{j-bins} Poiss[n_j^i | (K^i \lambda)_j + b_j^i]$$

$$\tilde{\chi}^2 = -2 \ln L(\lambda)$$



2 : 2

Regularisation

TUnfold

Tikhonov regularisation:

Combine

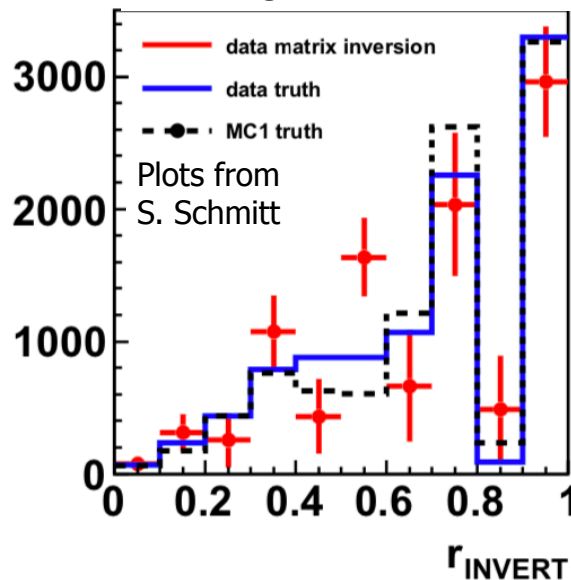
$$\tilde{\chi}^2 \rightarrow \tilde{\chi}^2 + \tau^2 ||\mathbf{L}(\lambda - \lambda_0)||^2$$

Bias vector

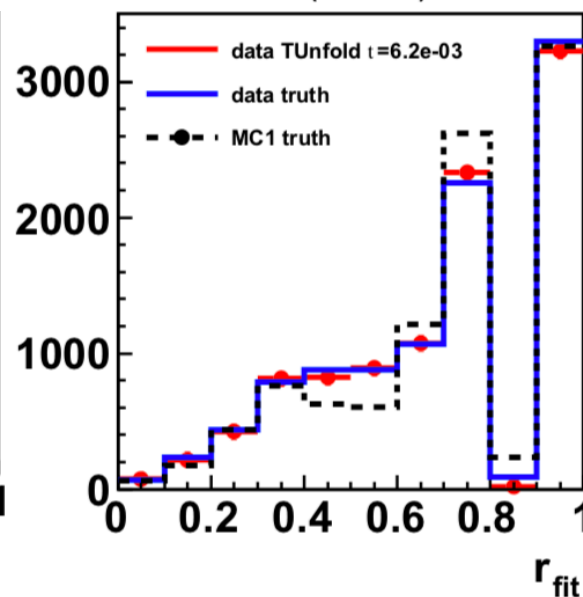
Regularisation strength

Matrix: unity, 1st or 2nd derivatives

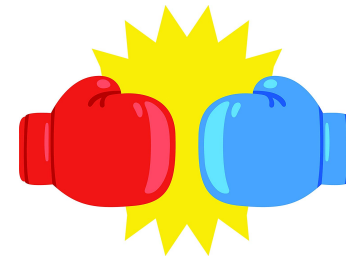
No regularisation



L-curve scan (TUnfold) nRec=100



Note: should use objective criteria to choose tau value, Here L-curve scan



3 : 3

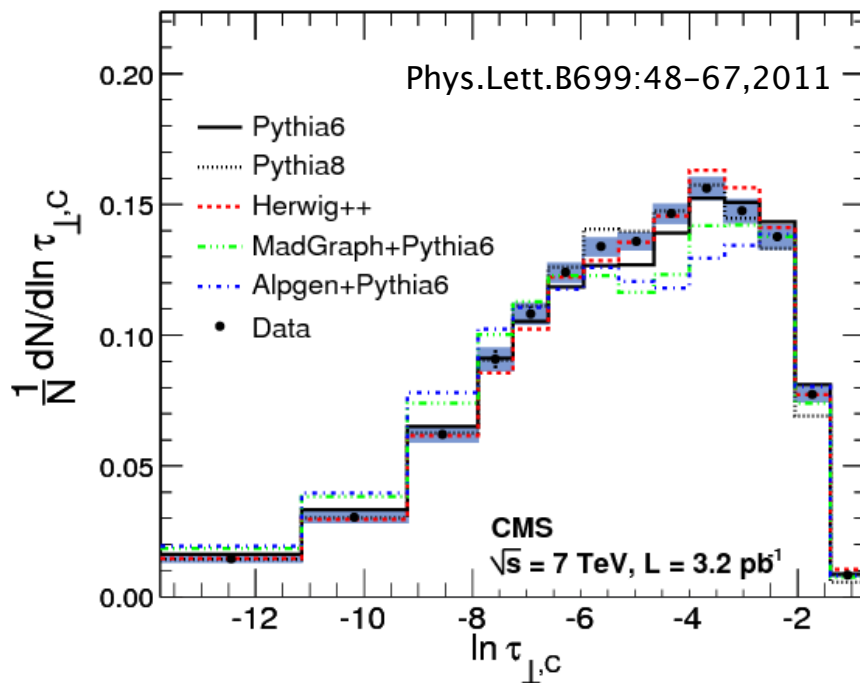
How can we make sure the constraint towards Bias vector (typically nominal signal MC prediction) is not too strong? → **bottomline test**

Bottomline test

Idea: Comparing theory models to unfolded data should NOT be more discriminative than comparing folded theories to detector data

→ assess with data-model χ^2 tests at both levels (stat. uncert. only)

Hadronic event shape analysis example with SVD unfolding



Material from talk M. Weber, PHYSTAT 2011

MC Generator	χ^2 values between data and smeared mc	χ^2 values between unfolded data and Gen mc
PYTHIA6	421	398
HERWIG++	211	200
MADGRAPH	2590	2570
ALPGEN	3860	3860

→ Chi2 order is the same before and after unfolding, values are similar

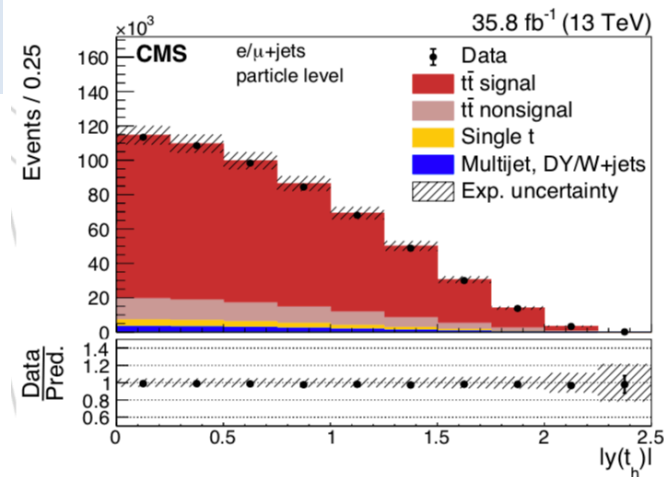
😊 No sign of too strong regularisation in this example

Bottomline test in CMS lepton+jets tt analysis

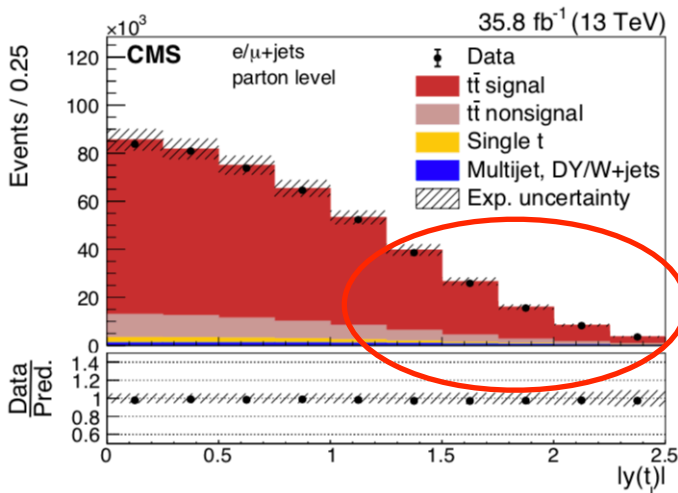
TOP-17-002
PRD 97, 112003 (2018)

D'Agostini
unfolding

$y(t)$ at det. level

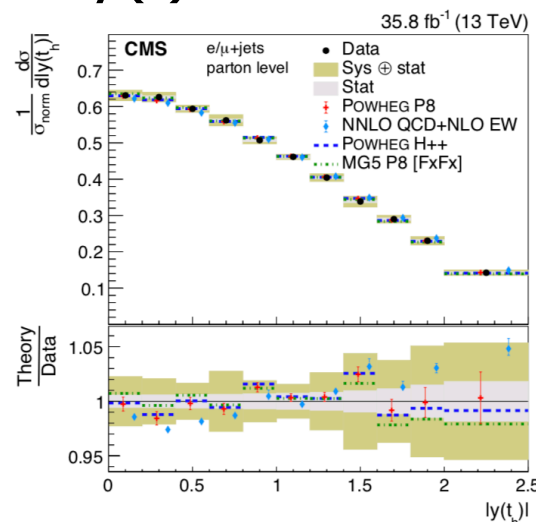


$t \rightarrow bj j'$

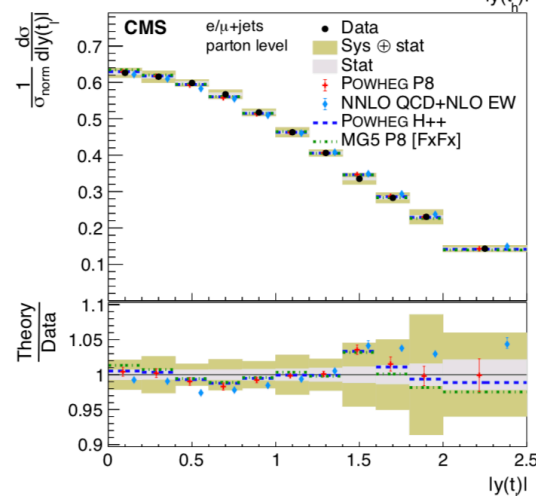


$t \rightarrow blv$

$y(t)$ unfolded to parton level



Need ~ 5 iterations
to have $\chi^2_{\text{data-models}}$
similar at both levels



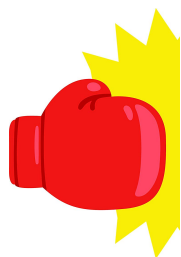
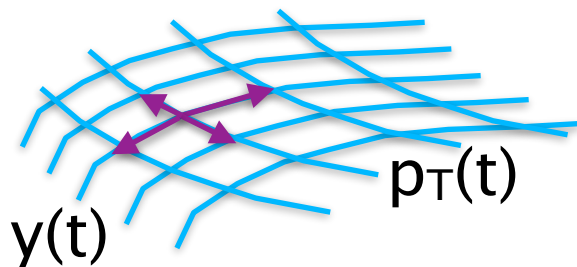
Need ~ 50 iterations
to have $\chi^2_{\text{data-models}}$
similar at both levels

- Number of needed iterations depend on bin-size/resolution
- Don't use blindly RooUnfold D'Agostini unfolding default of 4 iterations

Multi-differential cross section capabilities

TUnfold

- Automated mapping of multi-D input to internal 1D vectors
- Multi-D Tikhonov regularisation



4 : 3

Combine

- Multi-D unfolding performed e.g. in HIG-17-025, $H \rightarrow \gamma\gamma$

$$\frac{d^2\sigma}{dp_T^{\gamma\gamma} dN_{jet}}$$

w/o regularisation

- So far max. total number of unfolded bins $\mathcal{O}(15)$

Background treatment

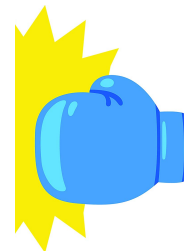
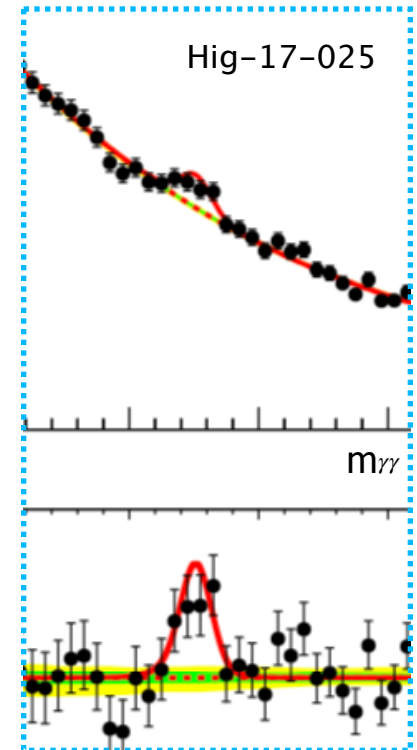
TUnfold

- User can supply for each background source:
 - bin-wise contribution + uncertainty
 - global normalisation factor + uncertainty
- TUnfold subtract backgrounds and propagate errors

4 : 4

Combine

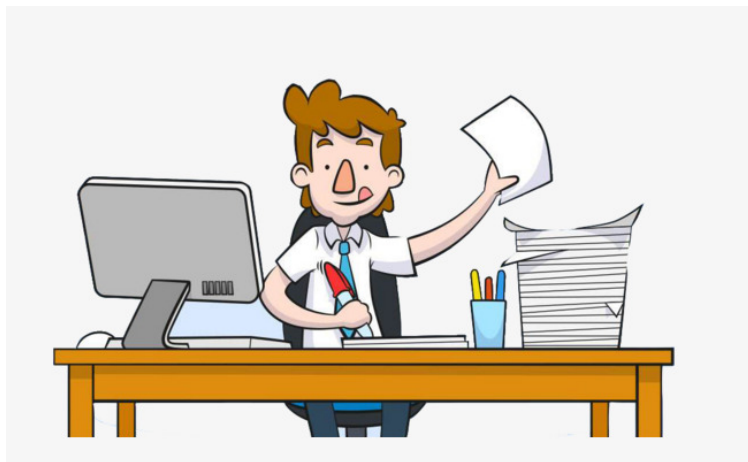
- Signals and backgrounds can be constrained from additional event categories, e.g. fitting to $m_{\gamma\gamma}$ spectrum



Systematic uncertainties

TUnfold

- Signal detector correction systematics \rightarrow varied response matrices
- Nominal matrix K , $\pm 1\sigma$ variations for source m : K_{m+} , K_{m-}
- Unfold with $K_{m+} \rightarrow$ cross section variation $\delta \hat{\lambda}_{m+}$
- TUnfold user has to perform all variations for all sources m and add to covariance matrix of λ

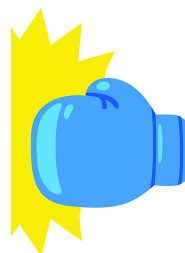


Combine

- Nuisance par θ_m , morph $K(\theta_m)$ from K_{m-} , K and K_{m+}
$$L \sim \prod_{j-\text{bins}} \text{Pois}[n_j | (K(\theta)\lambda)_j + b_j]$$

$$L \rightarrow L \cdot \text{Pdfs}(\theta)$$
- Perform profile likelihood fit, data can also constrain θ

4 : 5



MC Statistical uncertainties

TUnfold

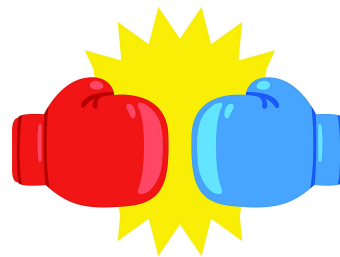
Combine

Statistical uncertainties of the response matrix K are:

- Error propagated to $V_{\hat{\lambda}}$

- Barlow Beeston lite for each element of K , \rightarrow nuisance parameter with gaussian PDF

John Conway, arXiv1103.0354



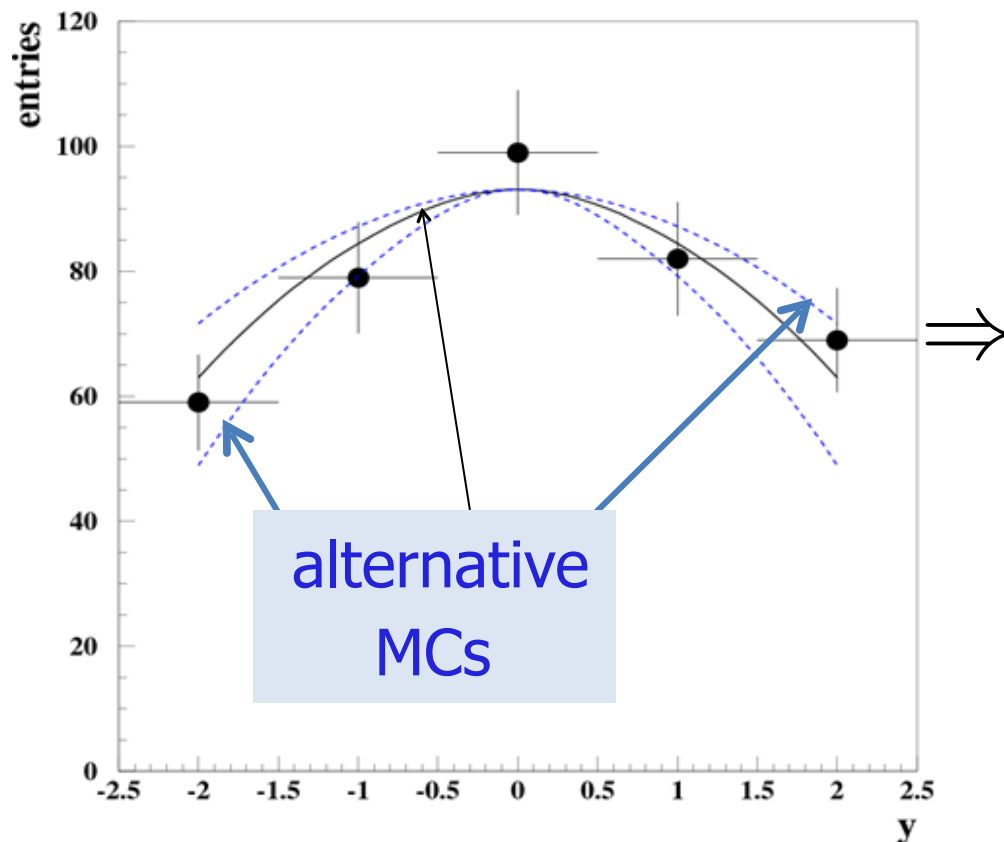
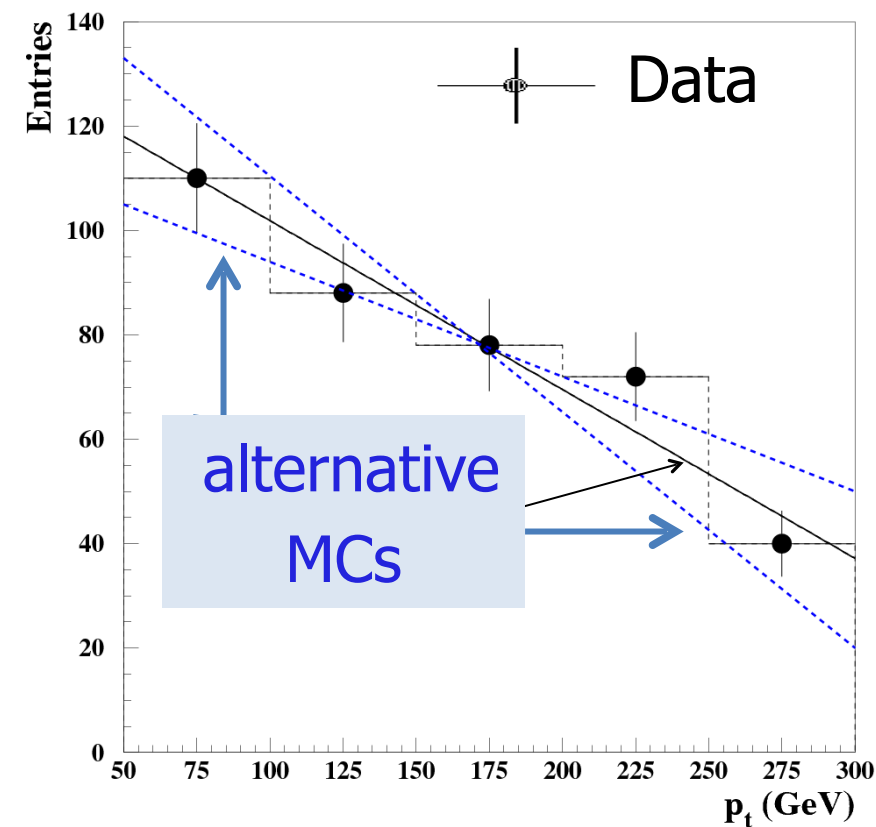
5 : 6



shutterstock.com • 305688461

Model systematics

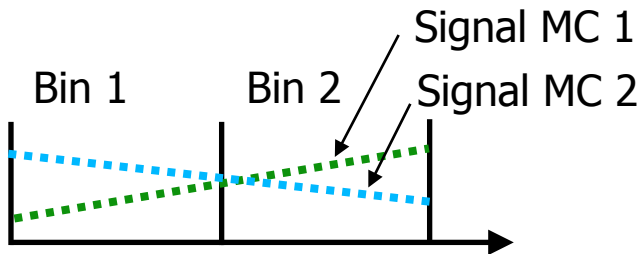
MC generators model physics imperfectly
→ systematics of response matrix



- Recommend simultaneous multi-D unfolding of p_T , y and other distributions and *projecting* results to 1D spectra of p_T , y , etc.
- Reduces model systematics (assumptions on the distribution of the variables integrated over)

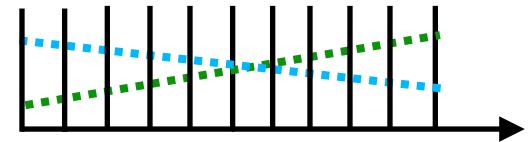
Wide bins and model systematics

M. Kuusela

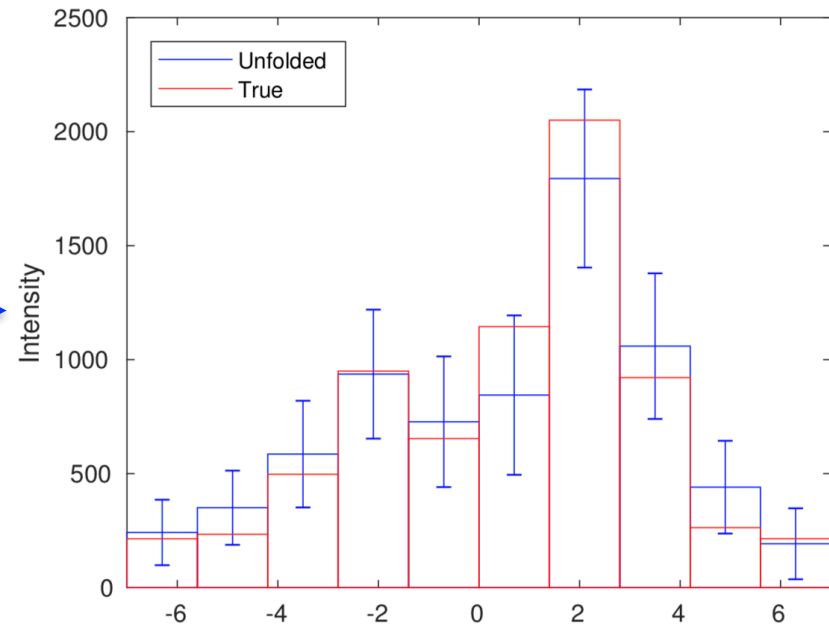
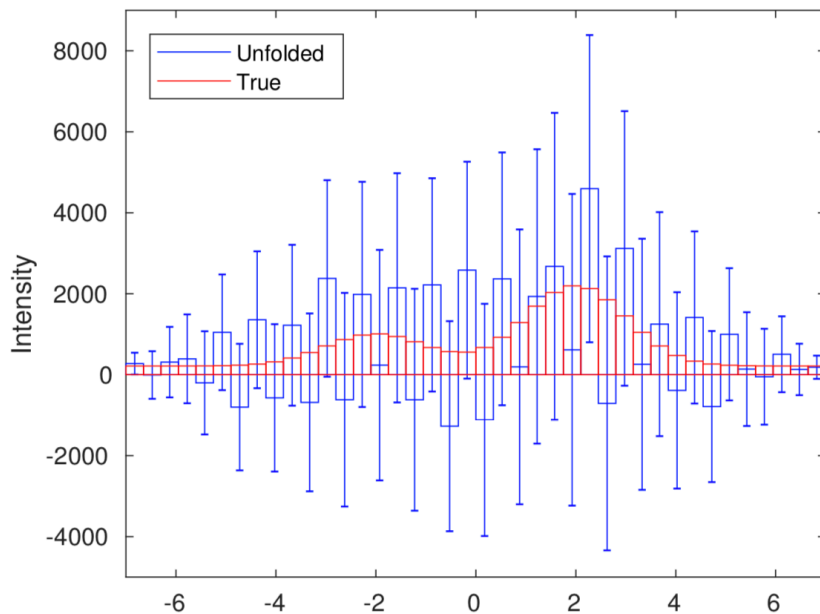


Different distributions within bins
→ Different response matrices
→ Model Systematics

Reduced effects within smaller bins



- Idea: unfold with fine bins and then aggregate results in wide bins



→ Improved modelling systematics, as detailed in talk

<https://indico.desy.de/indico/event/22731/session/5/contribution/24/material/slides/0.pdf>

Summary

TUnfold ideal for analyses:

- high statistics
- low background
- Multi-differential $d\sigma$
- Example: Top-pair $d\sigma$

Combine ideal for analyses:

- Any statistics, data can constrain background and systematics
- First $d\sigma$ for low statistics, small s/b
- High precision measurements
- Example: $d\sigma$ for single Top, Higgs

Not discussed tool:

- SVD similar than TUnfold but less flexible



Unfolding is not a simple black box: requires thoughts and checks on optimal choices for binning, regularisation, systematics, needed MC statistics, and many other things

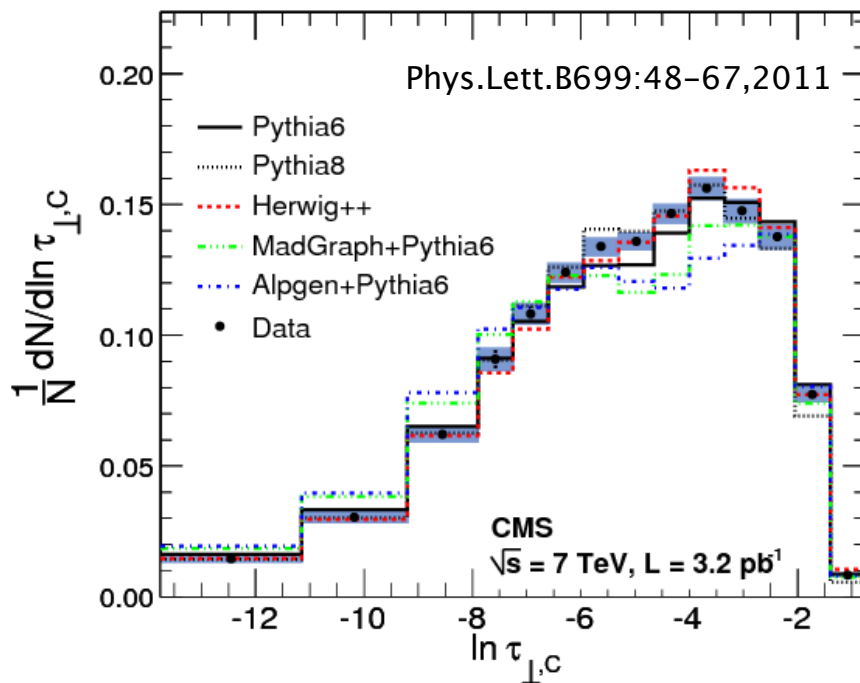
Backup slides

Bottomline test

Idea: Comparing theory models to unfolded data should NOT be more discriminative than comparing folded theories to detector data

→ assess with data-model χ^2 tests at both levels (stat. uncert. only)

Hadronic event shape analysis example with SVD unfolding



Material from talk M. Weber, PHYSTAT 2011

MC Generator	χ^2 values between data and smeared mc	χ^2 values between unfolded data and Gen mc
PYTHIA6	421	398
HERWIG++	211	200
MADGRAPH	2590	2570
ALPGEN	3860	3860

→ Chi2 order is the same before and after unfolding, values are similar



No sign of too strong regularisation in this example