

# Discussions on unfolding problems, methods and solutions – a physics analysis perspective

Bogdan Malaescu  
(LPNHE Paris - CNRS, CERN)

*bogdan.malaescu@cern.ch*



17/05/2018



# Content

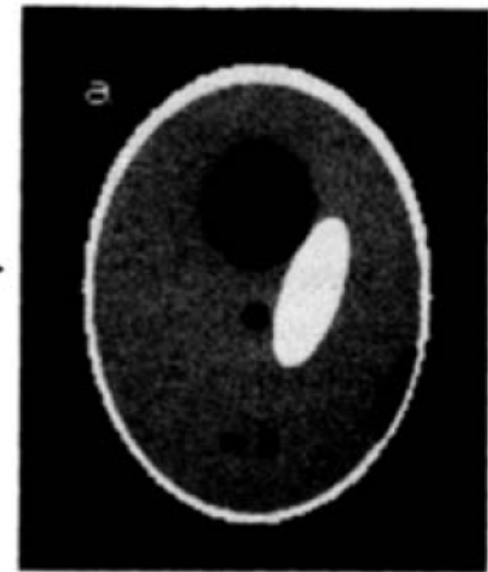
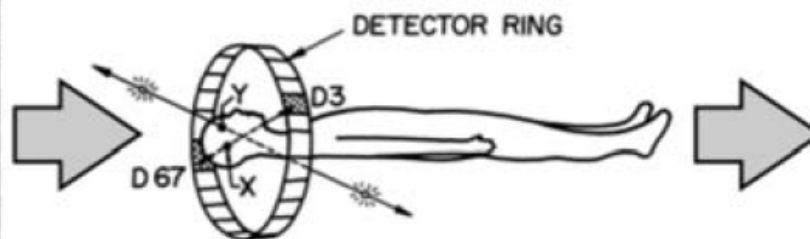
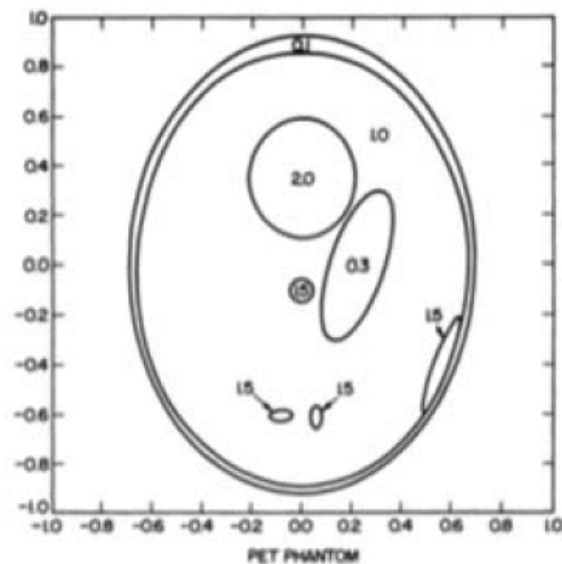
- Introduction: “ What is the (un)folding about ”
- Choice of phase-space
- Unfolding methods
- Methods for parameter setting, uncertainty evaluation and propagation
- How do we publish our data: HEPData entries for unfolded results
- Physics studies using unfolded data
- Multidimensional unfolding + Choice of binning

ATLAS (internal) recommendations:

<https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/StandardModelUnfoldingNew>

<https://cds.cern.ch/record/1694351>

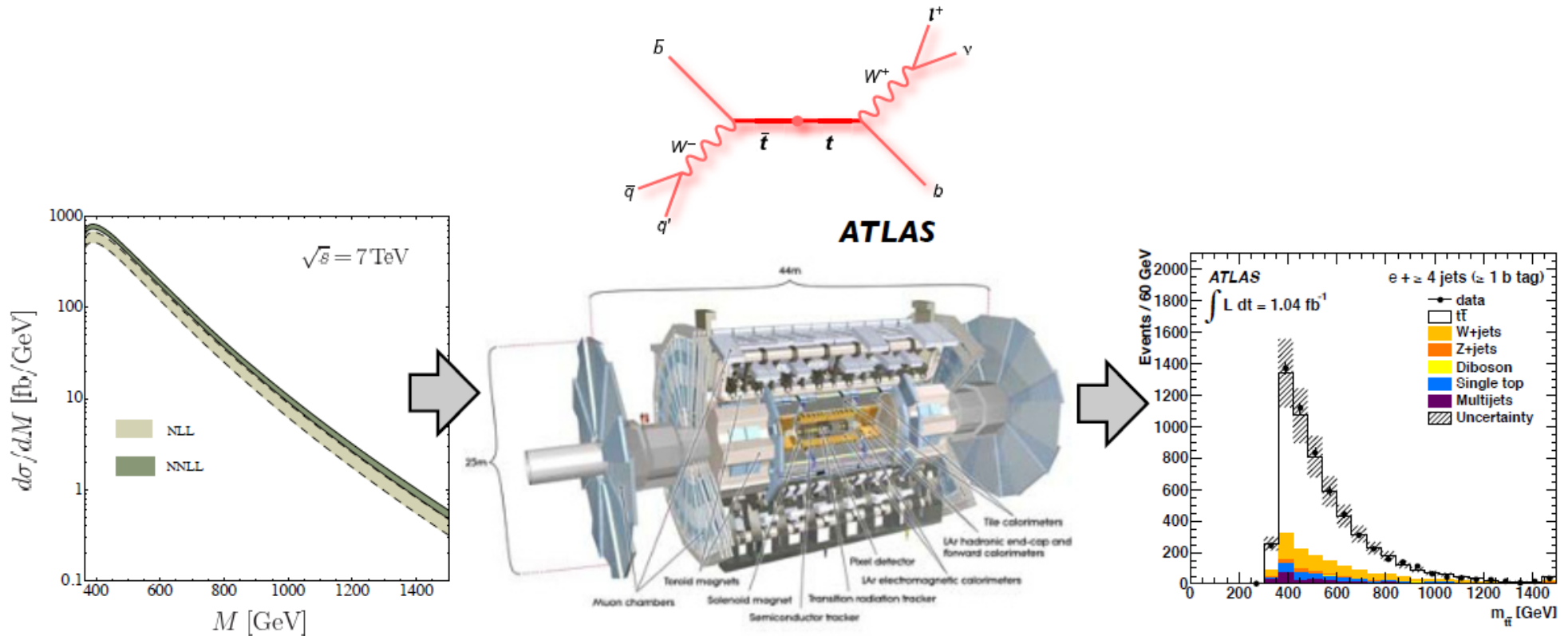
# Examples of (un)folding problems: PET scan



## Positron Emission Tomography

Y. Vardi et al.  
<http://www.jstor.org/stable/2288030>

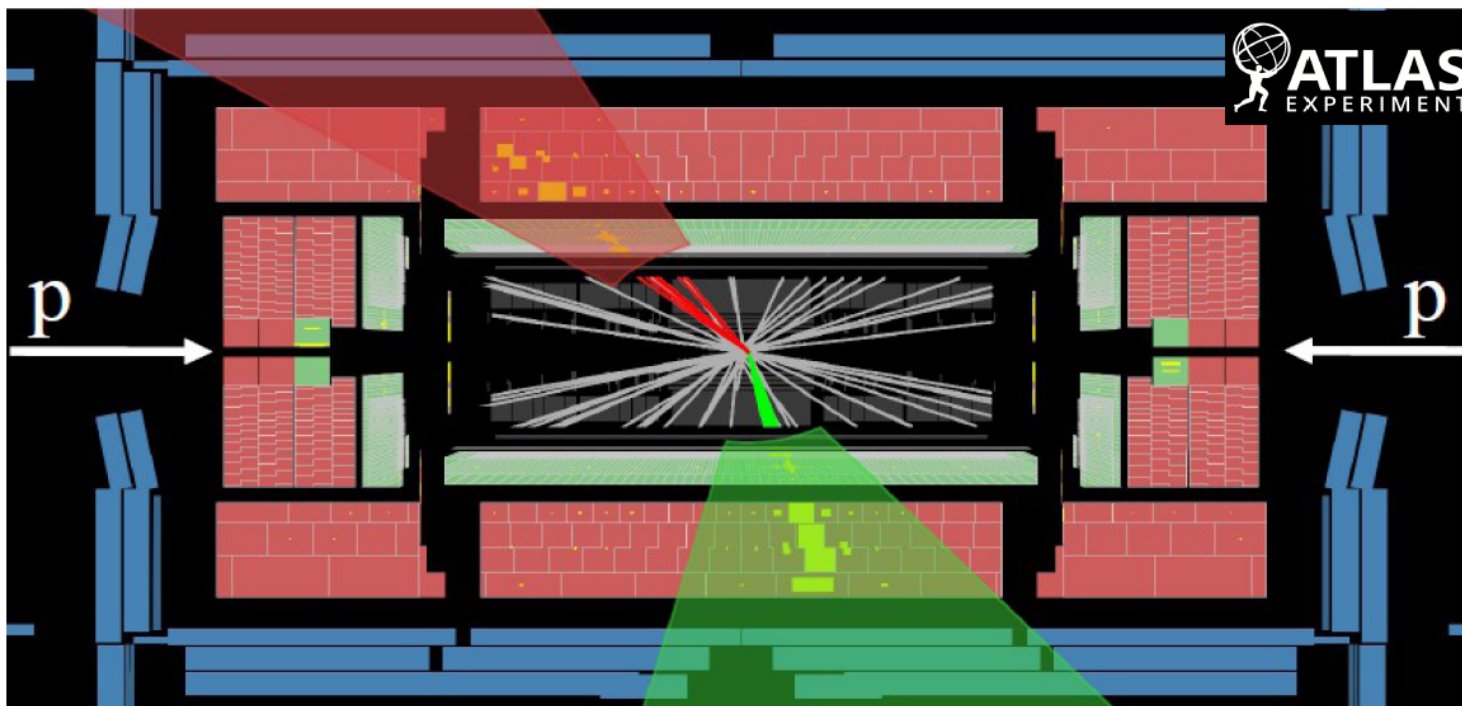
# Examples of (un)folding problems: top pairs @ LHC



F. Spano

EPJ Web of Conferences 55, 03002 (2013)

# Examples of (un)folding problems: jets @ LHC

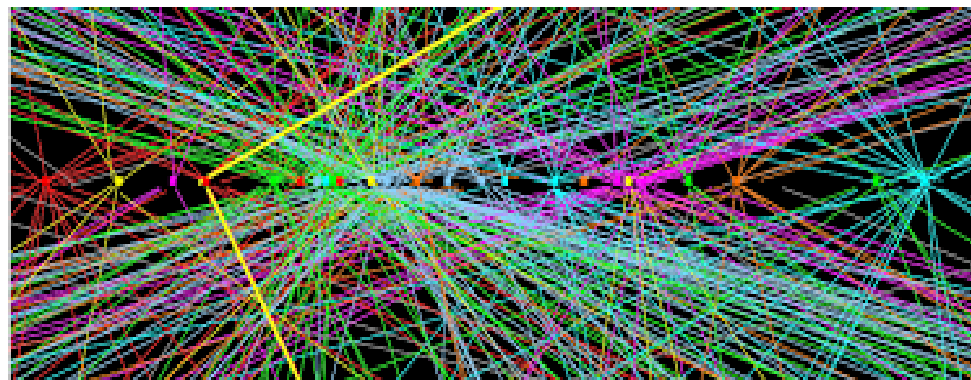
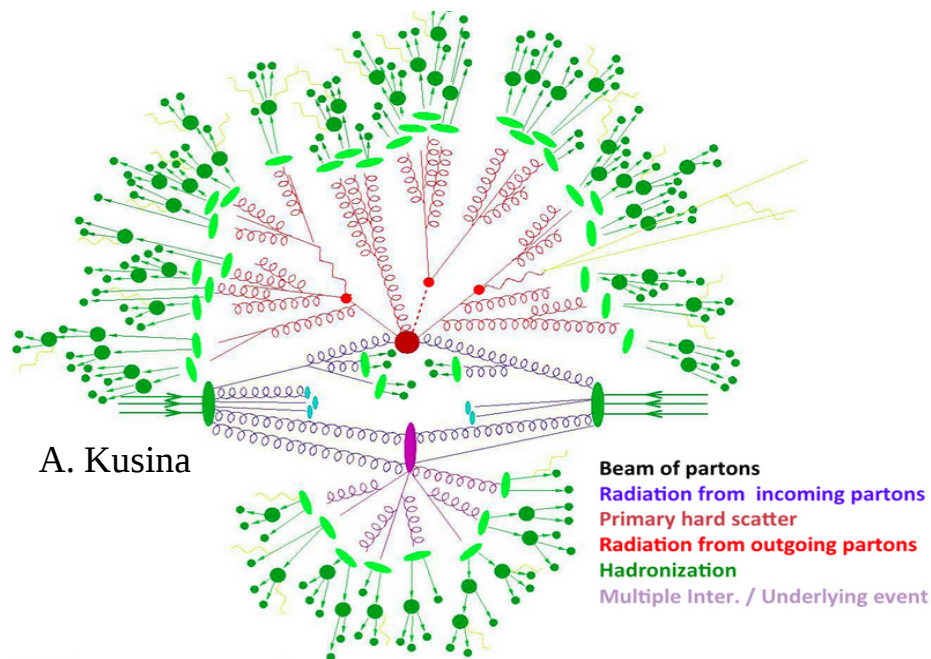


**Jets:** “sprays” of (quasi-)stable particles called hadrons, observed in the ATLAS detector  
→ **proxy to fundamental interactions in Nature:** Explore the limits of SM

# Environment and strategy of the jet studies

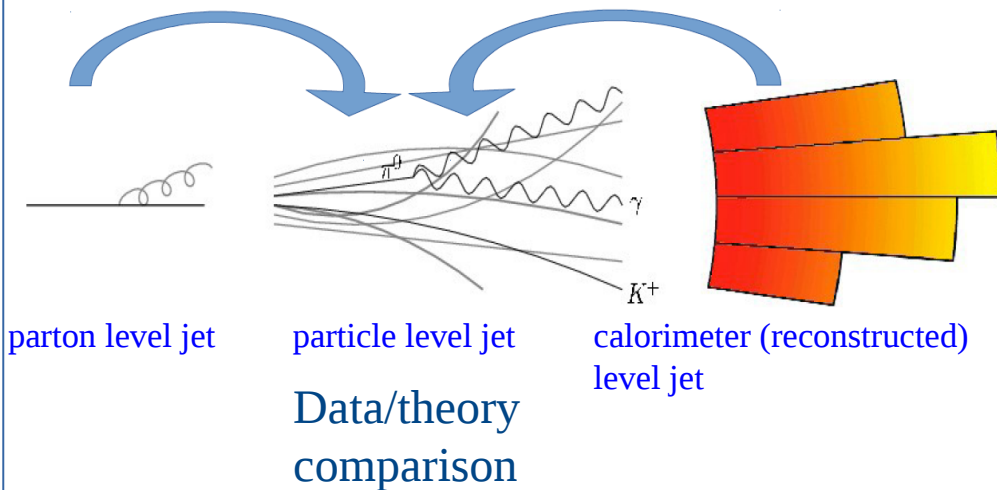
Typical proton-proton collision: a complex process in a difficult environment

Pile-up



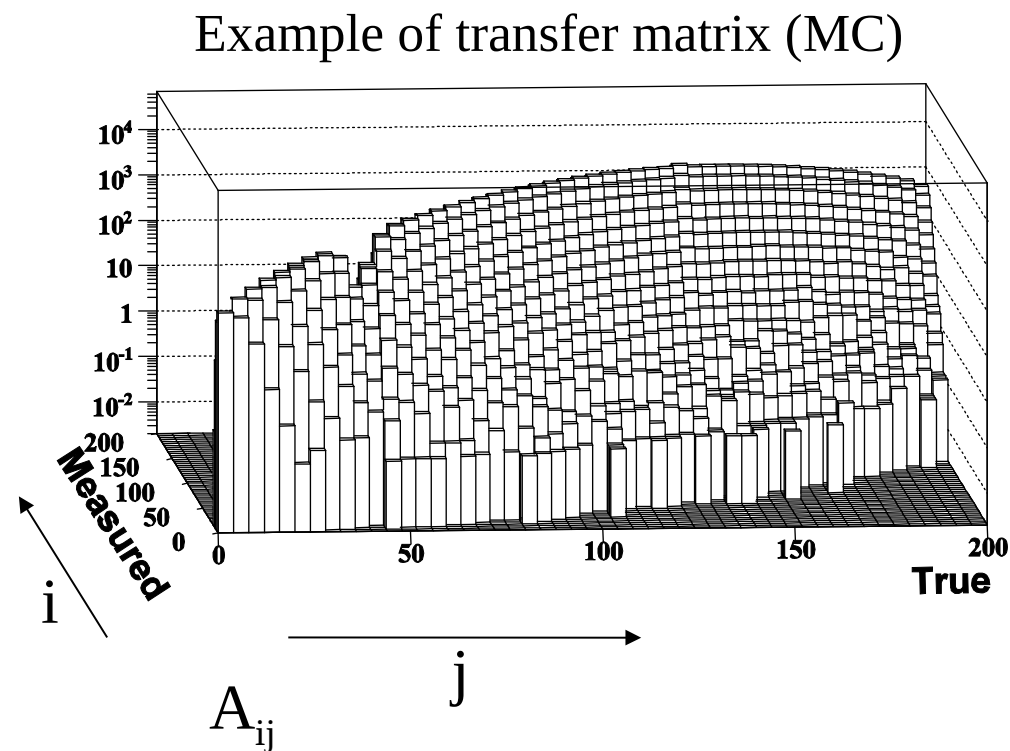
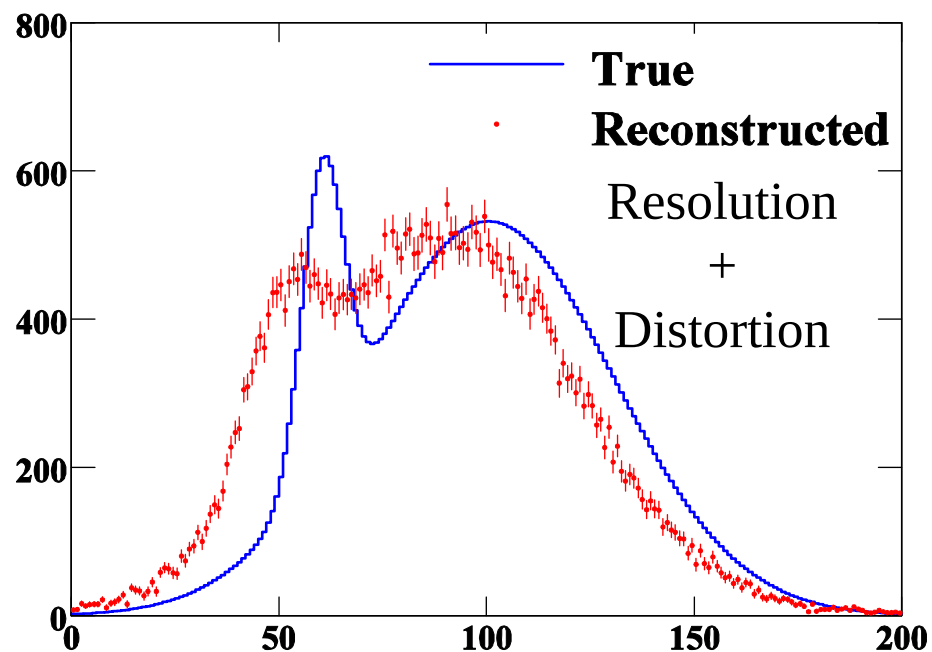
NP corrections  
 Hadronization & UE

Calibration+Unfolding  
 Jet energy response & resolution



Goal: *publish data corrected for detector effects (with minimal bias and minimal model dependence), with the full information needed for comparisons with theory predictions*

# Detector effects, folding and unfolding



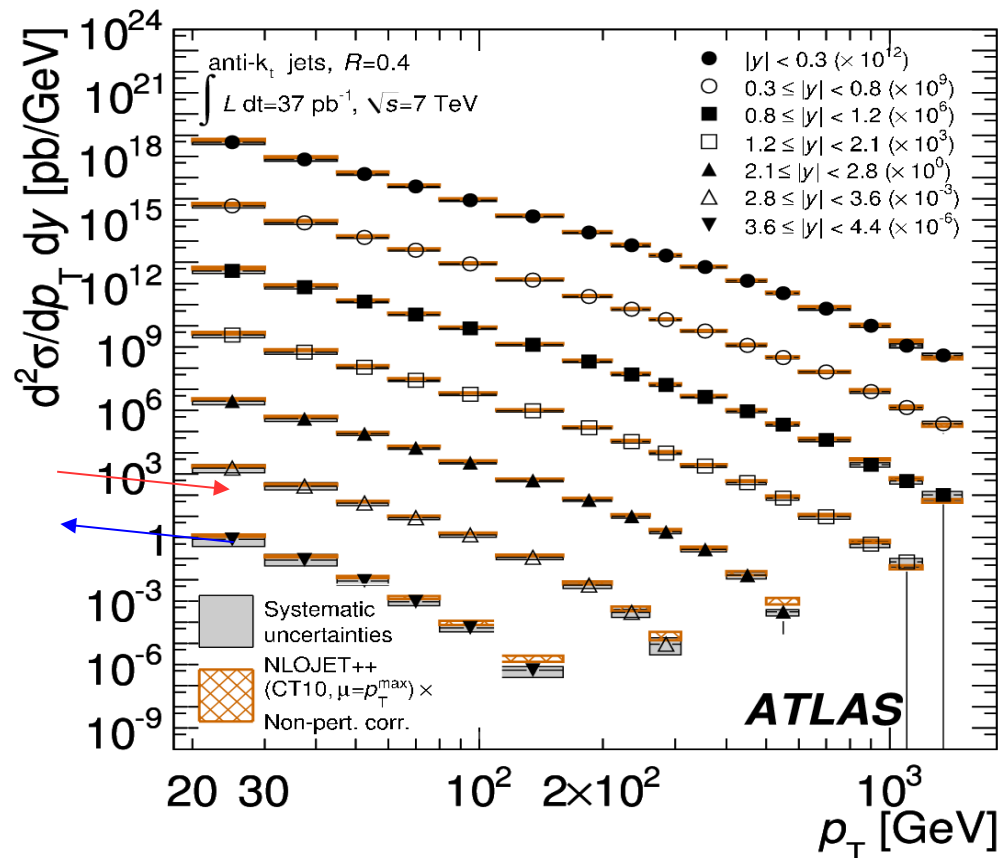
- Folding:  $f_{\text{meas}}(x) = \int R(x|y) f_{\text{true}}(y) dy$

$$P_{ij} = \frac{A_{ij}}{\sum_{k=1}^{n_d} A_{kj}} ; d = P \cdot t$$

- Focus on unfolding of detector effects (acceptance correction factorized)
- Unfolding is not a simple numerical problem  
→ regularization methods necessary

# Choice of the phase-space

- Selection defining phase-space at “truth” level – as close as possible to the reconstructed-level selection: *minimize extrapolation to reduce model dependence*
- Include over-/under-flow bins when migrations to the region of interest are relevant → These extra bins are generally not published



[arXiv:1112.6297](https://arxiv.org/abs/1112.6297)



# Unfolding methods + choice of one method & settings

- Maximum likelihood / matrix inversion
  - SVD ( + Tikhonov regularization )
  - Iterative Bayes-inspired regularized unfolding
  - Full Bayesian unfolding
  - Iterative, dynamically stabilized (IDS) method
  - Bin-by-bin correction :  $d_i \cdot (T_i/R_i)^{MC}$  → potentially large bias by relying on truth MC (used only when small bin-to-bin migrations & for statistics limited measurements e.g. Higgs differential Xsec; cross-check with matrix-based method)
  - In general, recommended not to (dis)favor some particular method
  - Recommended to evaluate the performance of *several methods & regularizations* and use the “optimal” one for the given use-case
- Take into account: *systematic uncertainty related to the unfolding method (bias due to MC/data shape difference & regularization)*; impact on statistical uncertainties & correlations; constraints induced on binning choice

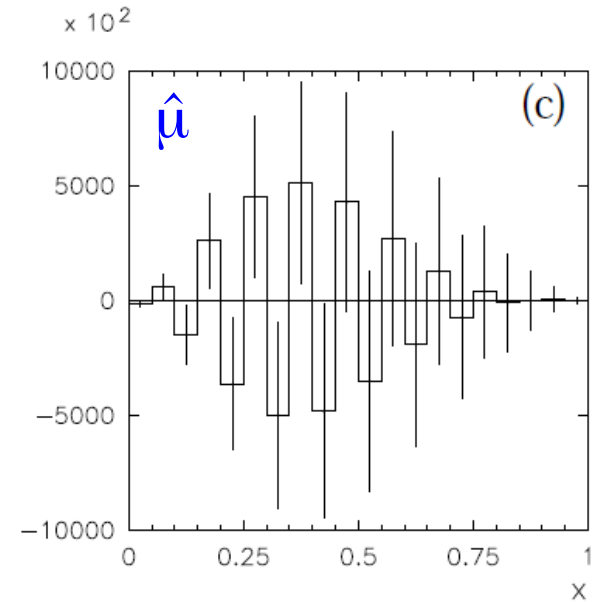
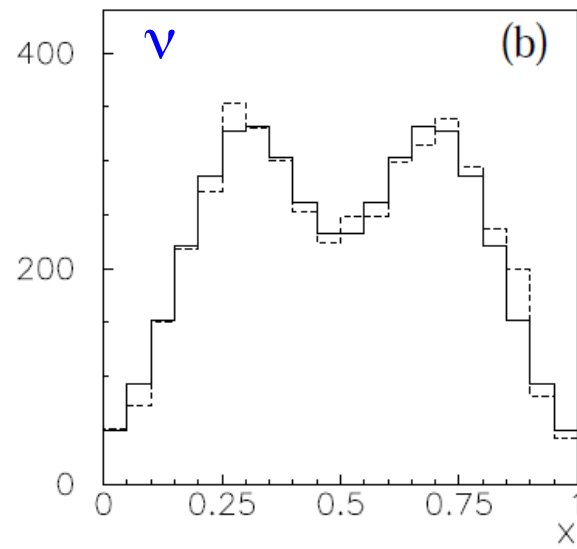
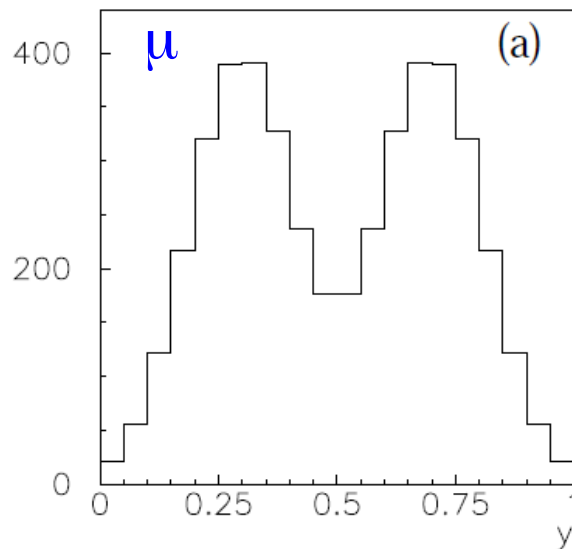
# Matrix inversion

Folding of signal and background in data:

$$E[\mathbf{n}] = \boldsymbol{\nu} = R\boldsymbol{\mu} + \boldsymbol{\beta}$$

Unfolding based on matrix inversion:

$$\hat{\boldsymbol{\mu}} = R^{-1}(\mathbf{n} - \boldsymbol{\beta})$$



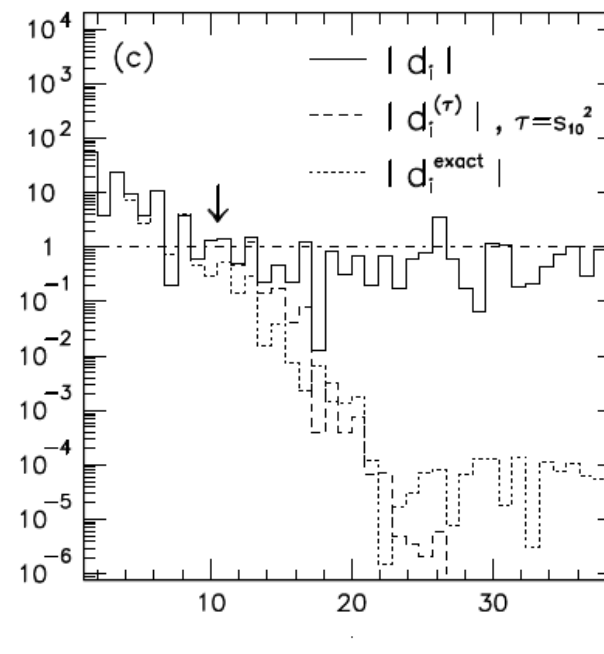
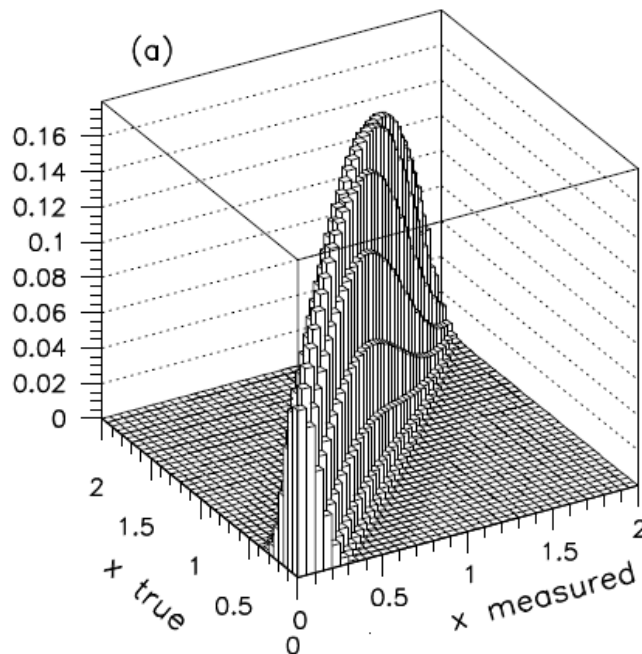
→ Result unbiased, but with large variances

# Singular Value Decomposition (SVD)

→ Inspired by the matrix inversion, but with regularization:

Suppress effect of small eigenvalues ( $\sim$ noise) + constraint on smoothness of the unfolded distribution → Regularization (may introduce bias)

$$S(\boldsymbol{\mu}) = - \sum_{i=1}^{M-2} [(\mu_{i+2} - \mu_{i+1}) - (\mu_{i+1} - \mu_i)]^2$$



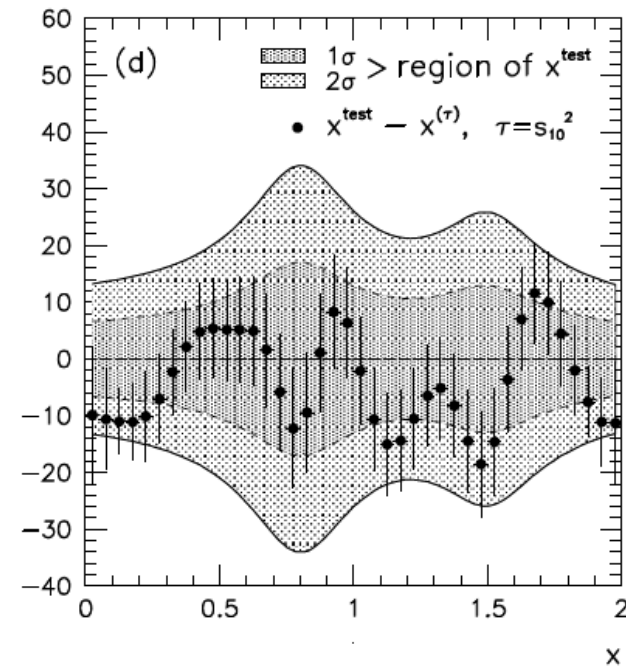
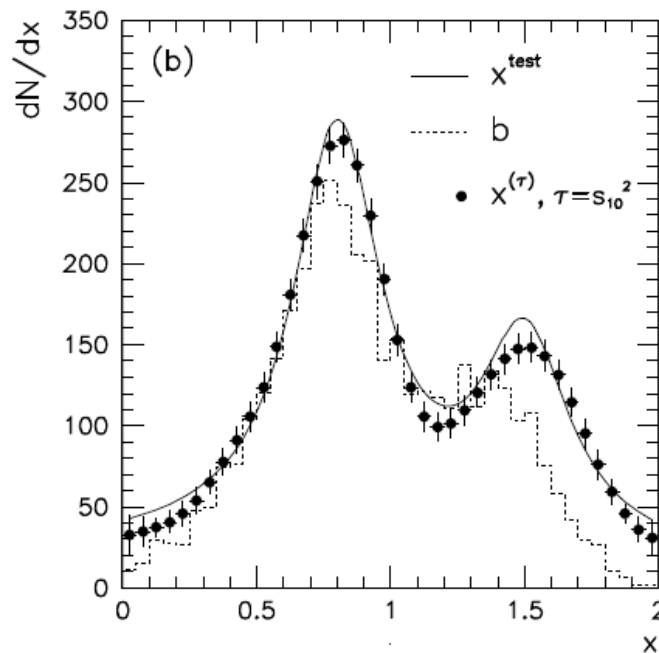
Nucl. Instr. Meth. A 372, 1996 (469)

# Singular Value Decomposition (SVD)

→ Inspired by the matrix inversion, but with regularization:

Suppress effect of small eigenvalues ( $\sim$ noise) + constraint on smoothness of the unfolded distribution → Regularization (may introduce bias)

$$S(\boldsymbol{\mu}) = - \sum_{i=1}^{M-2} [(\mu_{i+2} - \mu_{i+1}) - (\mu_{i+1} - \mu_i)]^2$$



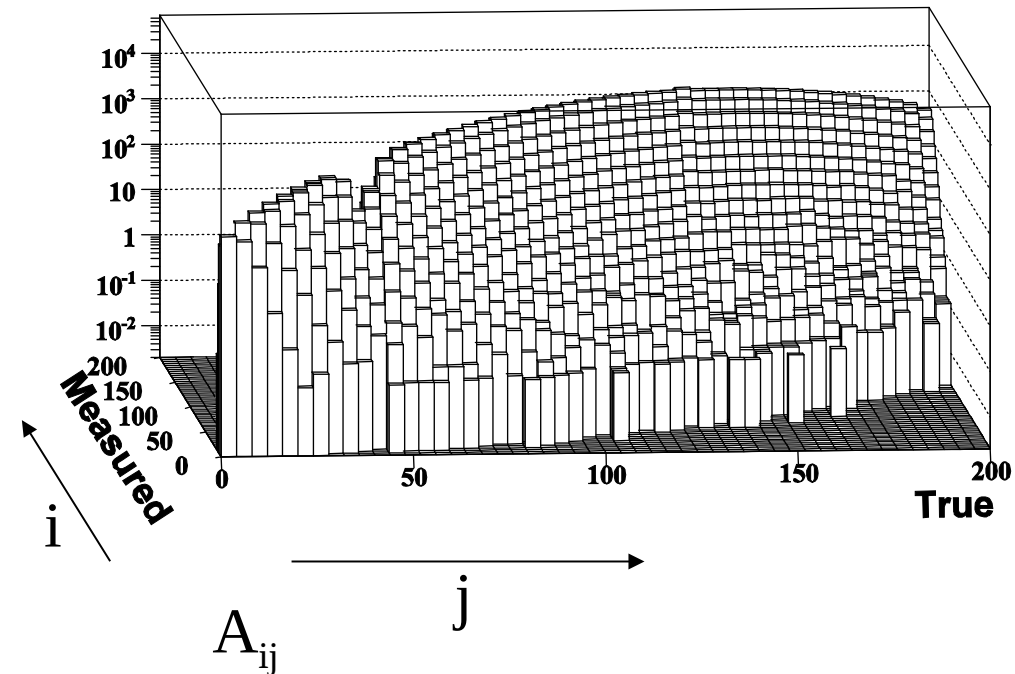
Nucl. Instr. Meth. A 372, 1996 (469)

# An Iterative, Bayes-inspired Unfolding Method

$$P_{ij} = \frac{A_{ij}}{\sum_{k=1}^{n_d} A_{kj}}$$

$$\tilde{P}_{ij} = \frac{A_{ij}}{\sum_{k=1}^{n_u} A_{ik}} ; u = \tilde{P} \cdot d$$

→ Note:  $\tilde{P}_{ij}$  depends on the shape of the truth distribution in MC



- 1<sup>st</sup> unfolding, where the original transfer matrix is used

→ 1) Transfer matrix improvement (hence of the unfolding probability matrix)

Reweight the truth MC distribution based on previous unfolding result.

→ 2) Improved unfolding

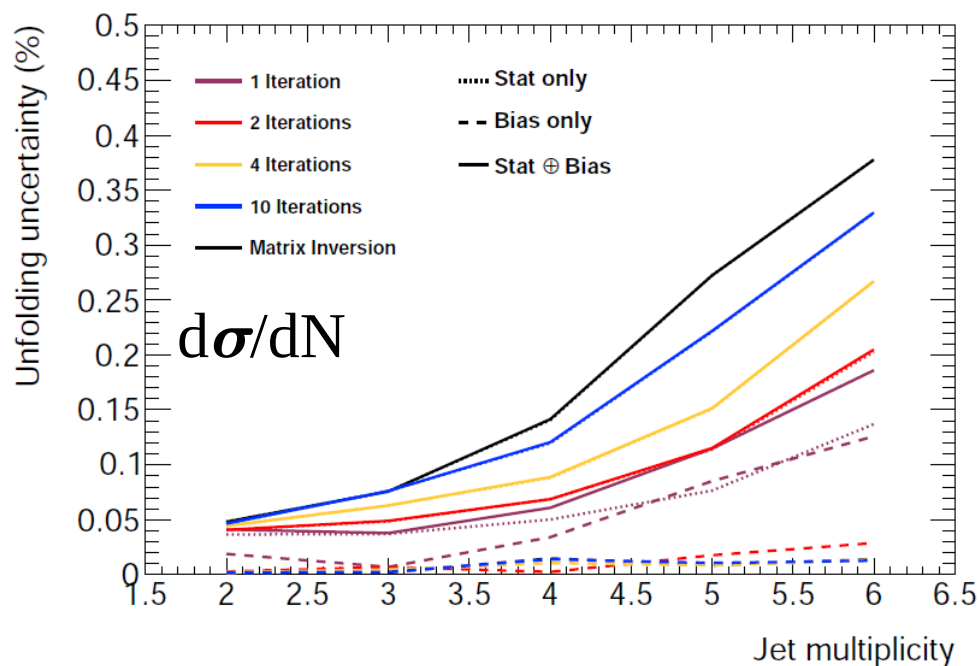
→ Choice on number of iterations = regularization

→ Other methods exist, like e.g. dynamical regularization from the treatment of fluctuations in each bin, at each step of the procedure

# Iterative methods: choice of the number of iterations

- Number of iterations = regularization parameter (the corresponding recommendations apply)

Dustin Henry Urbaniec's PhD



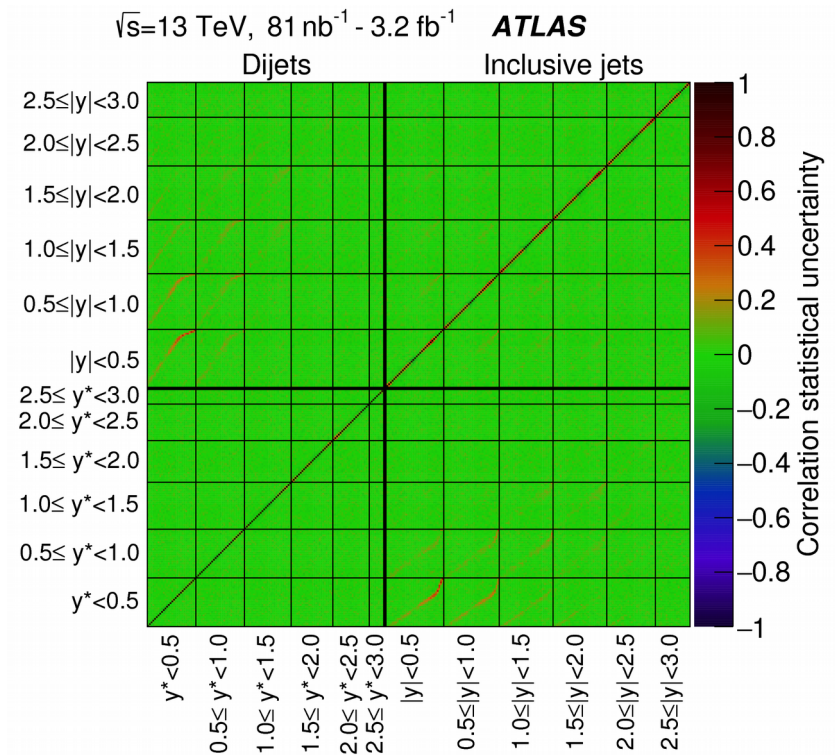
- Compare data and the modified reconstructed MC: see how much information is left to be propagated from the data shape to the truth MC shape  $\rightarrow$  bin-by-bin comparison or using a  $\chi^2$
- Suggestion by d'Agostini: compare results from consecutive steps  $\rightarrow$  risk of  $\sim$ small changes between consecutive steps, while having a significant bias

# Statistical uncertainties

- Due to both data and MC
- Propagated using pseudo-experiments done separately/simultaneously for data and MC

## → Bootstrap method

- multiply event weights  
by random number:  $\text{Poisson}(1)$
- seed given by event number
- allows to correlate measurements  
with overlapping samples

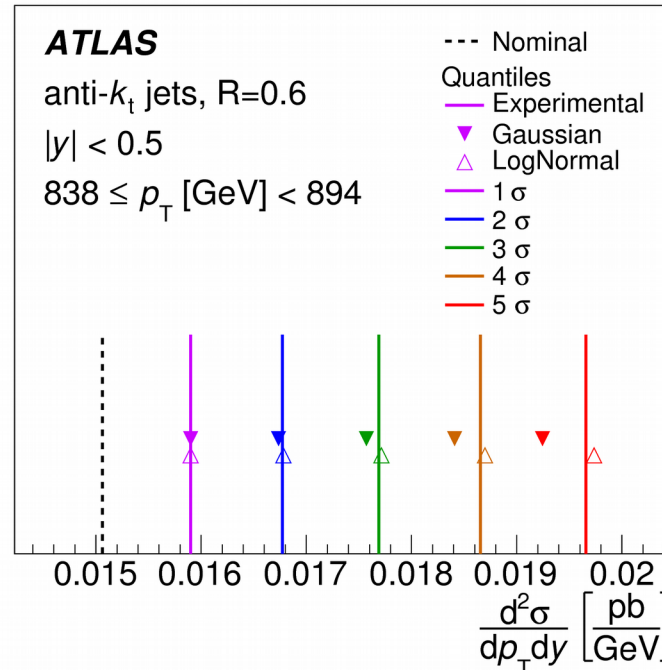


- Publish covariance matrix and/or a series of results based on each pseudo-experiment (i.e. Bootstrap replicas)
- Some unfolding methods provide estimates of the stat uncertainties

## → recommend cross-check with pseudo-experiments

# Propagation of systematic uncertainties from inputs

- Modify input (pseudo-)data spectrum by  $\pm 1\sigma$  of the uncertainty, re-do unfolding and compare with nominal result
- can also use 1...5 $\sigma$  scans or pseudo-experiments



arXiv:1410.8857

- can shift reconstructed spectrum in transfer matrix instead of input spectrum: switched positive and negative variations
- For resolution uncertainties, perform smearing of the transfer matrix: smearing factor given by quadratic difference between resolution enhanced by  $1\sigma$  and nominal resolution



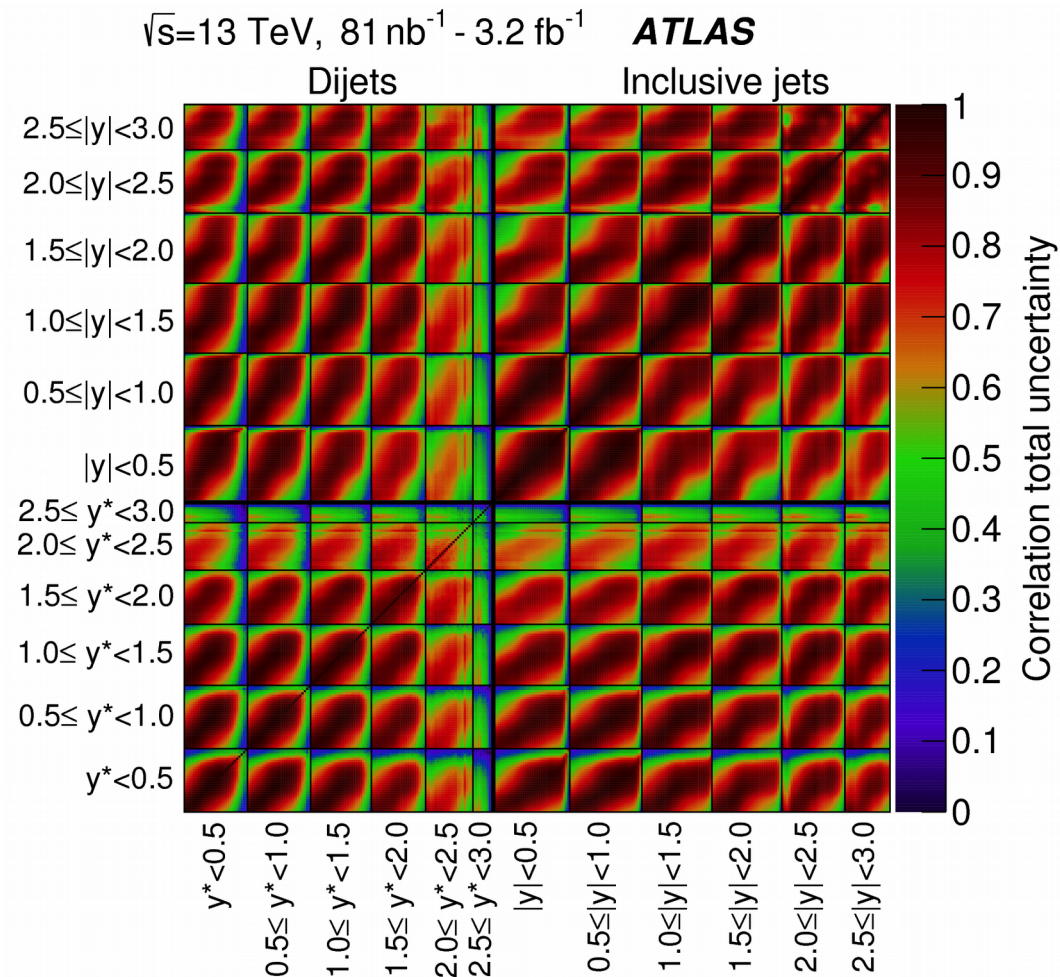
# Propagation of systematic uncertainties from inputs

- Bootstrap method to evaluate statistical uncertainties on the propagated systematics + rebinning/smoothing; [arXiv:1312.3524](#)
- Alternative propagation using pseudo-experiments (more difficult to probe e.g.  $5\sigma$  effects)
- Alternative propagation option: include uncertainties as nuisance parameters in the definition of the response matrix + profile likelihood or Bayesian marginalization (often used for folding/template fits)

# Propagation of systematic uncertainties from inputs

- Split of systematics in sub-components (fully correlated in phase-space, independent between each-other) allows to evaluate correlations between different phase-space regions and between different measurements
- Information made available in HEPData

$$Cov_{ij} = \sum_{k=1}^{N_{syst}} S_i^k \cdot S_j^k$$

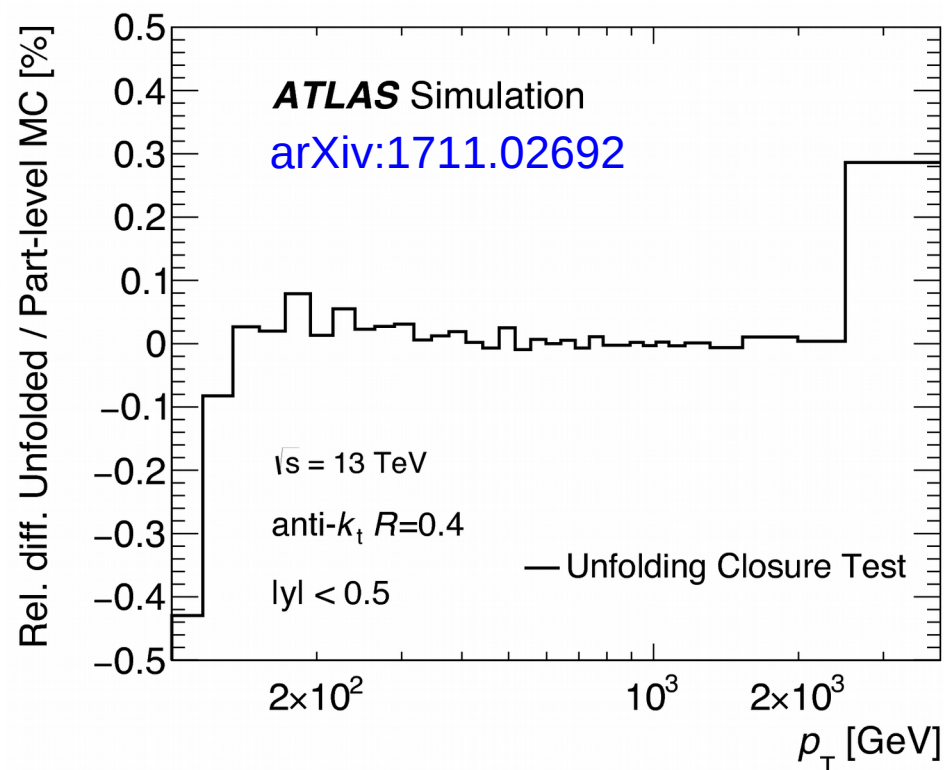
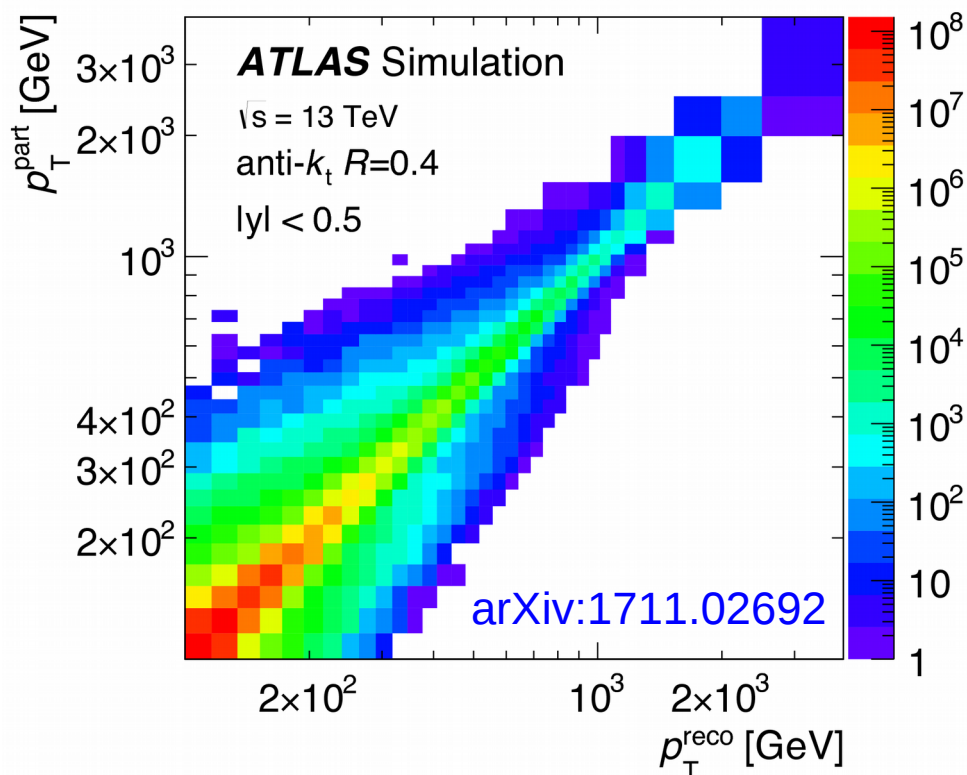


# Tests of the unfolding

- “Technical closure test” → same MC for the transfer matrix and input distribution (pseudo-data)
- “Data-driven closure test” → allows to evaluate a systematic related to the unfolding method and the choice of regularization (see next slides)
- “Linearity test” → MC samples with various truth inputs; check linear dependence between unfolded and truth values of a quantity of interest
- “Pull test” → relevant only for unfolding methods providing an estimate of the stat uncertainty (i.e. not from pseudo-experiments)

# Data-driven closure test<sup>(\*)</sup>: motivation, procedure, example

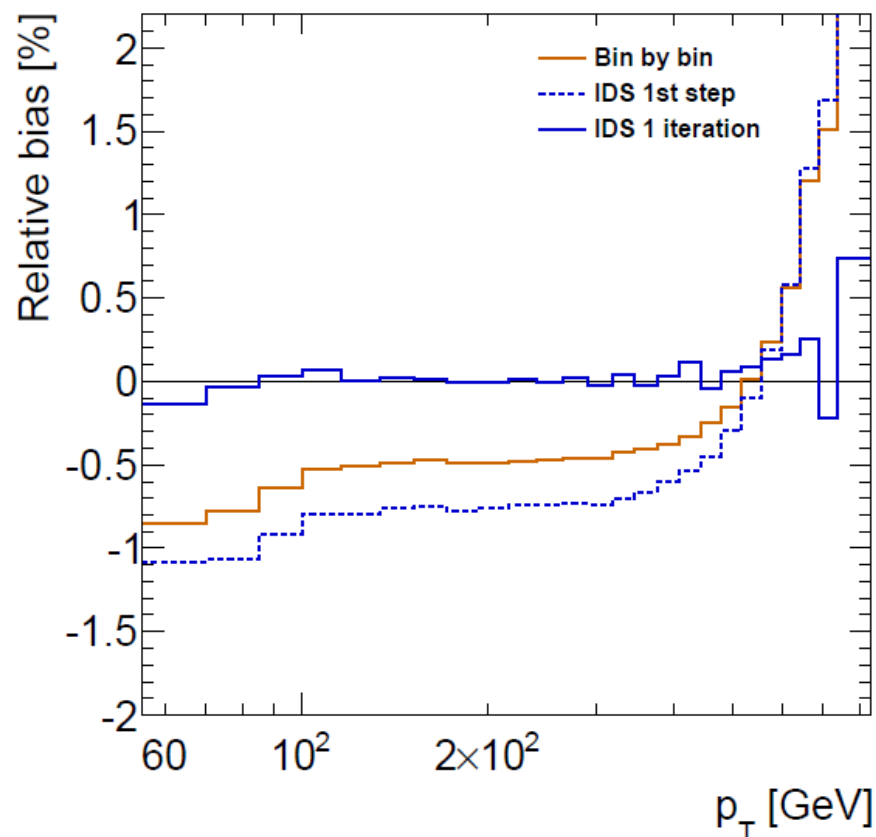
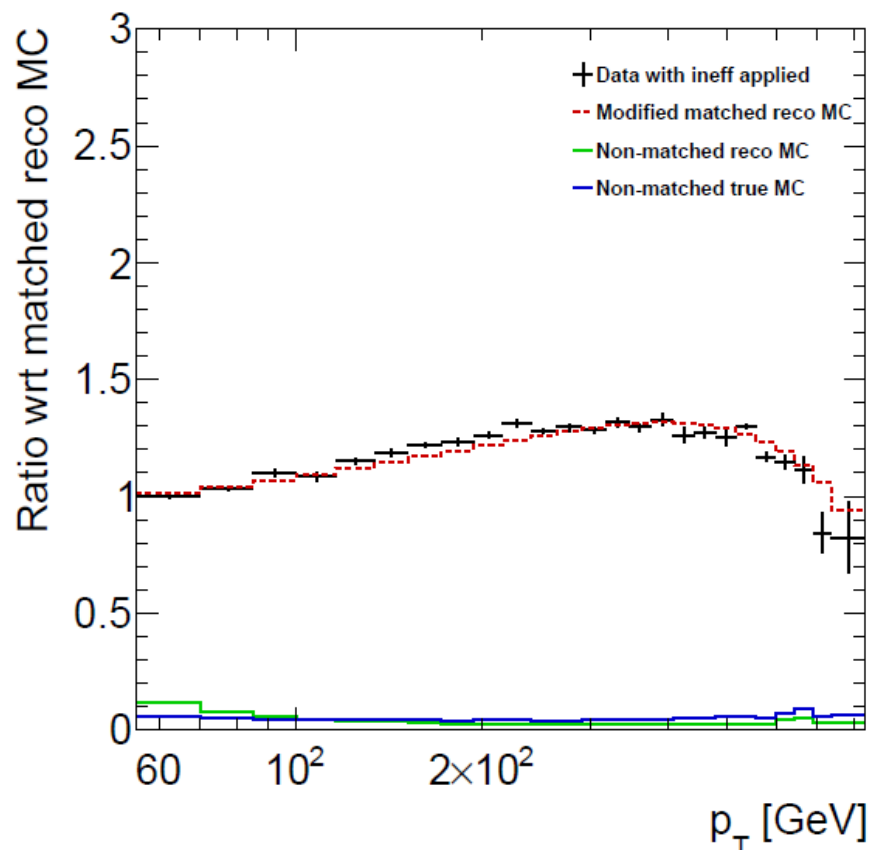
- In-situ determination of the **unfolding uncertainty** related to the MC shape (i.e. to the **data/MC shape difference**) and to the **regularization** :
  - reweight true MC by smooth function: improved data/recoMC agreement;
  - unfold the reweighted reconstructed MC;
  - compare with reweighted true MC.



(\*) Method introduced in arXiv:0907.3791, used in arXiv:1112.6297 etc.

# Data-driven closure test<sup>(\*)</sup>: motivation, procedure, example

- In-situ determination of the **unfolding uncertainty** related to the MC shape (i.e. to the **data/MC shape difference**) and to the **regularization** :
  - reweight true MC by smooth function: improved data/recoMC agreement;
  - unfold the reweighted reconstructed MC;
  - compare with reweighted true MC.



(\*) Method introduced in [arXiv:0907.3791](https://arxiv.org/abs/0907.3791), used in [arXiv:1112.6297](https://arxiv.org/abs/1112.6297) etc.

# Some subtleties

# Data-driven closure test: remarks

- In general, when the data/MC shape differences increase, the unfolding bias increases
- Still, in general, the unfolding bias  $\ll$  shape differences ( see previous slides )
- Folding of modified truth MC can be done with nominal/stat fluctuated resolution matrix: the 2<sup>nd</sup> option requires a series of toys – median converges towards nominal
  - Data/MC shape comparison: is some shape difference “worth” being taken into account in the reweighting?
    - in general, binned data/MC comparison taking into account stat uncertainties
    - one can perform more quantitative comparisons (e.g. through  $\chi^2$  evaluations) taking into account data and MC systematics
      - shape comparison: the MC normalization free parameter in the  $\chi^2$  evaluations
      - has to be done globally and more locally (for restricted phase-space regions): in global comparison, local systematic differences can be “hidden” by the ndof++

# *Analyses with several “correction steps”*

- Analyses generally involve a series of corrections for: efficiency, acceptance, migrations, subtraction of background / "fakes" etc.
- Guidelines:
  - coherent selection criteria applied to the data and the reconstructed MC samples (used to derive corrections), at each correction step;
  - coherence between data and the reconstructed MC for what concerns the variables as a function of which the corrections are derived / applied;
  - coherence between the motivation / derivation / application of the corrections;
  - minimize the model dependence for all the corrections.



# Background subtraction; Propagation of systematics

- **Background subtraction** (data-driven or estimated with MC )

→ (Generally) **performed before unfolding**, because resolution is generally different for signal and background;

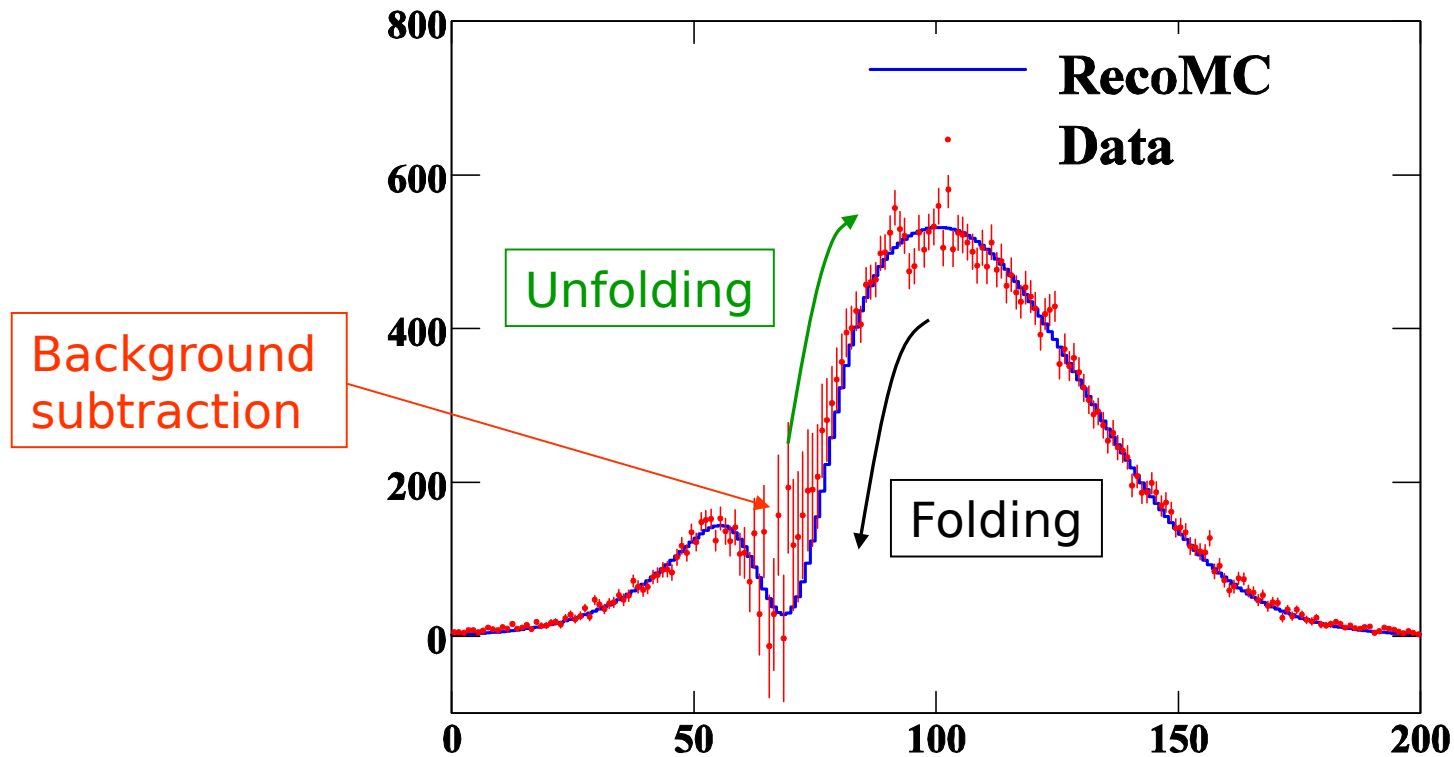
(otherwise, special care is required when building the transfer matrix)

- **Propagation of the systematic uncertainties** (for the calibration(s) these often concern data-MC differences) from the input objects to the final analysis results: *avoid double-counting and/or fake cancellations*

→ **propagate systematics either through (pseudo-)data** (for input distributions, data-driven background estimates etc.) or **MC** (transfer matrix, MC-templates etc.), but not both;

→ **keep track of the signs** when evaluating and propagating the impact of a systematic variation on various inputs.

# Possible difficulties: fluctuations due to background subtraction



- A “standard” unfolding could propagate large fluctuations into precise regions of the spectrum
- Can be addressed by taking into account the uncertainties of the data points in the unfolding (used to compute the significance of data-MC differences in each bin - IDS)

# Treatment of “hidden variables”

- Check the impact on the average resolution, when integrating differently over various phase-space regions (e.g. central-forward regions when unfolding pT distribution)
- Does not concern the closure test, for which this extra reweighting is not needed (probes impact of data/MC shape difference for the observable of interest)
- A first check can be done by comparing unfolding results when using different MC samples (different truth-level distributions of variables other than the unfolded one)
  - truth MC shapes for the variable(s) of interest (i.e. the quantities that are unfolded) must be re-weighted such that they match between different MC samples: avoid double-counting with data-driven closure test
  - if significant differences are seen: compare with relevant resolution uncertainties and identify variables causing the difference
- Re-weight the transfer matrix as a function of more variables than the unfolded one and re-do the unfolding with modified matrix: amount of re-weighting should be data-driven

# How we report the results in HEPData

# W+jets @ 7TeV (2011) – HEPData information

<http://hepdata.cedar.ac.uk/view/ins1319490>

<b>ABS(ETARAP(LEPTON))</b> < 2.47 - (1.37 TO 1.52) <b>ABS(YRAP(JET))</b> < 4.4 <b>Cross section</b> $\sigma(W + N_{jets})$ [pb] <b>N_{jet}</b> 0.0 1.0 2.0 3.0 4.0 5.0 6.0 7.0 <b>PT(JET)</b> > 30 GeV <b>PT(LEPTON)</b> > 25 GeV <b>RE</b> P P --> W JET(S) X <b>SQRT(S)</b> 7000.0 GeV								
Source	Systematic uncertainty (In %)							
Jet energy scale EffectiveNP 1	-0.022	-1.6	-2.1	-2.7	-4.2	-7	-14	-9.1
Jet energy scale EffectiveNP 2	0.01	2.9	3.6	4.4	6.3	8.6	12	5.6
Jet energy scale EffectiveNP 3	-0.003	-1.2	-1.4	-1.7	-2.5	-3.3	-7	-4.8
Jet energy scale EffectiveNP 4	-0.005	0.038	0.027	0.074	0.33	0.52	3.2	2
Jet energy scale EffectiveNP 5	0.003	0.1	0.14	0.18	0.31	-0.53	-2.7	-3.2
Jet energy scale EffectiveNP 6	-0.001	-0.23	-0.29	-0.37	-0.61	-0.48	-0.49	-3.6
Jet energy scale EffectiveNP 7	-0.076	-5.1	-6.6	-9.5	-12	-15	-15	-12
Jet energy scale EffectiveNP 8	0.0	0.0	0.0	0.001	-0.006	-0.034	-0.12	-0.026
Jet energy scale EffectiveNP 9	-0.006	-1.1	-1.3	-1.8	-1.9	-1.2	-4.1	-2.5
Jet energy scale Pileup Offset Mu	-0.026	-0.62	-0.68	-1	-1.3	-2.1	-3.2	-5.3
Jet energy scale Pileup Offset NPV	-0.011	-0.23	-0.32	-0.4	-0.42	-2.1	-6	5.8

Split of systematic uncertainties in sub-components

...

# Inclusive jets @ 7TeV (2011) – HEPData information

<http://hepdata.cedar.ac.uk/view/ins1325553>

[Link to hepdata\\_tables\\_covM.tgz file containing statistical covariance matrices](#)

[Link to hepdata\\_tables\\_data\\_replicas.tgz file containing 10000 replicas \(used to derive statistical correlations between different ATLAS measurements\)](#)

[Link to hepdata\\_tables\\_NP and EW corrections.tgz file containing non-perturbative and electroweak corrections](#)

## Bootstrap replicas made public & used in e.g. 1602.01110

Table 1 ( Table 3. ) [HIDE DATA](#) or as: input, plain text, AIDA, PyROOT, YODA, ROOT, mpl, DMelt, MarcXML or YAML

Measured double-differential inclusive-jet cross section for the range  $0.0 \leq |y| < 0.5$  and for anti-kT jets with radius parameter  $R = 0.4$ . It is based on the data sample of proton-proton collisions at 7 TeV of centre-of-mass energy collected in 2011 by the ATLAS experiment at the LHC. The data sample corresponds to the integrated luminosity of  $4.5 \text{ fb}^{-1}$ .

The statistical uncertainties arising from data and MC simulation have been combined.

All the components of the systematic uncertainty are shown. They are: all the components of the jet energy scale uncertainty (jesX), the uncertainty of the jet energy resolution (jer), the uncertainty of the jet angular resolution (jar), the uncertainty of data unfolding (unfold), the uncertainty of the jet quality selection (qual), the luminosity uncertainty (lumi). All the components are assumed to be independent of each other. Each component is assumed to be fully correlated in pT and eta. Concerning the shape of the different components, Gaussian distribution assumption works for most of them. The three columns correspond to three different sets of the systematic uncertainty built with nominal, stronger or weaker assumptions on correlations between the jet energy scale uncertainty components. For more information on the systematic uncertainties, see the reference paper.

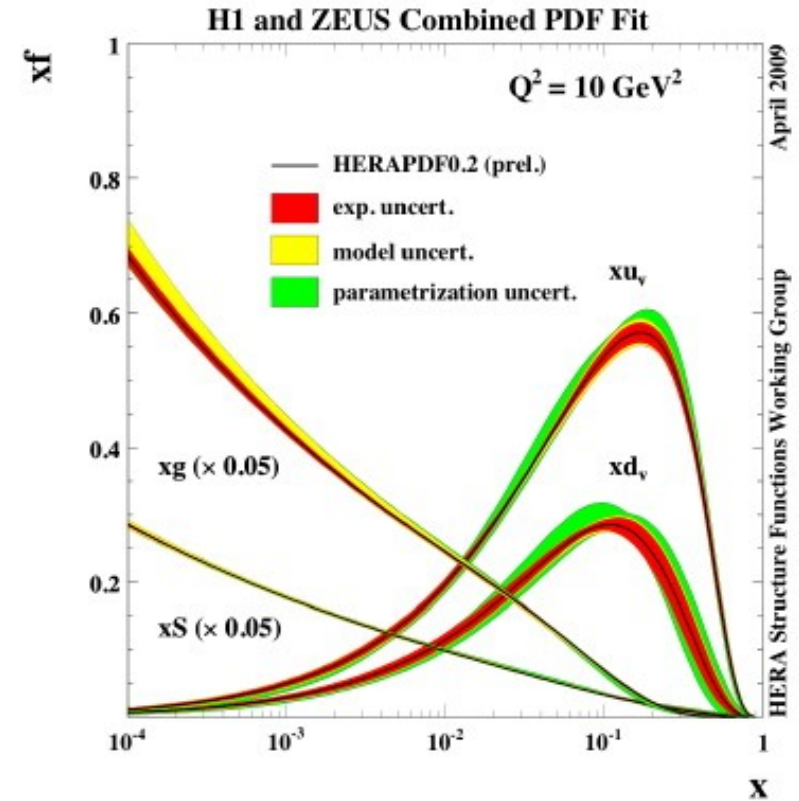
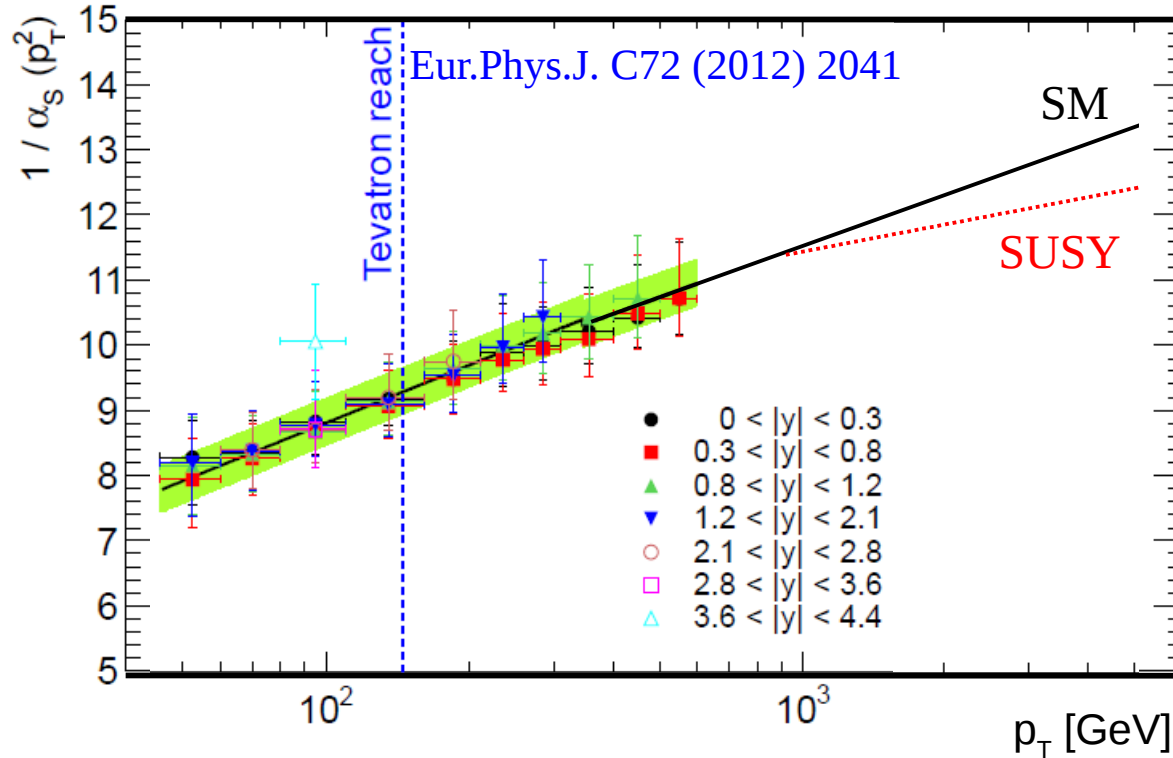
R	0.4
RE	P P --> JET X
SQRT(S)	7000.0 GeV
PT(JET) IN GEV	D2(SIG)/DPT(JET)/DYRAP IN PB*GEV**-1
100. - 116.	<p>4230 ± 0.89% (stat) ± 3.0% (sys,jes0)                      +0.4%,-0.5% (sys,jes1) ± 0.2% (sys,jes2)                      ± 0.3% (sys,jes3) ± 1.2% (sys,jes4)                      ± 1.0% (sys,jes5) +0.2%,-0.1% (sys,jes6)                      +0.5%,-0.6% (sys,jes7) ± 0.1% (sys,jes8)                      +0.8%,-0.9% (sys,jes9) ± 0.3% (sys,jes10)                      ± 0.8% (sys,jes11) +0.5%,-0.6% (sys,jes12)                      ± 0.00% (sys,jes13) ± 0.00% (sys,jes14)                      ± 0.00% (sys,jes15) ± 0.00% (sys,jes16)                      ± 0.00% (sys,jes17) ± 0.00% (sys,jes18)                      ± 0.00% (sys,jes19) ± 0.00% (sys,jes20)                      ± 0.1% (sys,jes21) +0.1%,-0.00% (sys,jes22)                      ± 0.00% (sys,jes23) ± 0.1% (sys,jes24)                      ± 0.1% (sys,jes25) ± 0.3% (sys,jes26)                      ± 0.4% (sys,jes27) +0.5%,-0.6% (sys,jes28)                      +0.4%,-0.3% (sys,jes29) ± 0.00% (sys,jes30)                      ± 0.00% (sys,jes31) ± 0.00% (sys,jes32)                      ± 0.00% (sys,jes33) ± 0.1% (sys,jes34)                      +0.00%,-0.1% (sys,jes35) +0.00%,-0.1% (sys,jes36)                      ± 0.00% (sys,jes37) ± 0.00% (sys,jes38)                      ± 0.00% (sys,jes39) ± 0.00% (sys,jes40)                      ± 0.00% (sys,jes41) ± 0.00% (sys,jes42)                      ± 0.00% (sys,jes43) ± 0.00% (sys,jes44)                      ± 0.00% (sys,jes45) ± 0.00% (sys,jes46)                      ± 0.00% (sys,jes47) ± 0.00% (sys,jes48)</p>

Split of systematic uncertainties in sub-components

## How we use the results

- PDF and/or  $\alpha_s$  fits
- Quantitative data / theory comparisons
- Limits on New Physics contributions
- Etc.

# Extraction of Physics information from measurements



→ Involves use of test statistics exploiting information on uncertainties and their correlations (between various measurement bins)

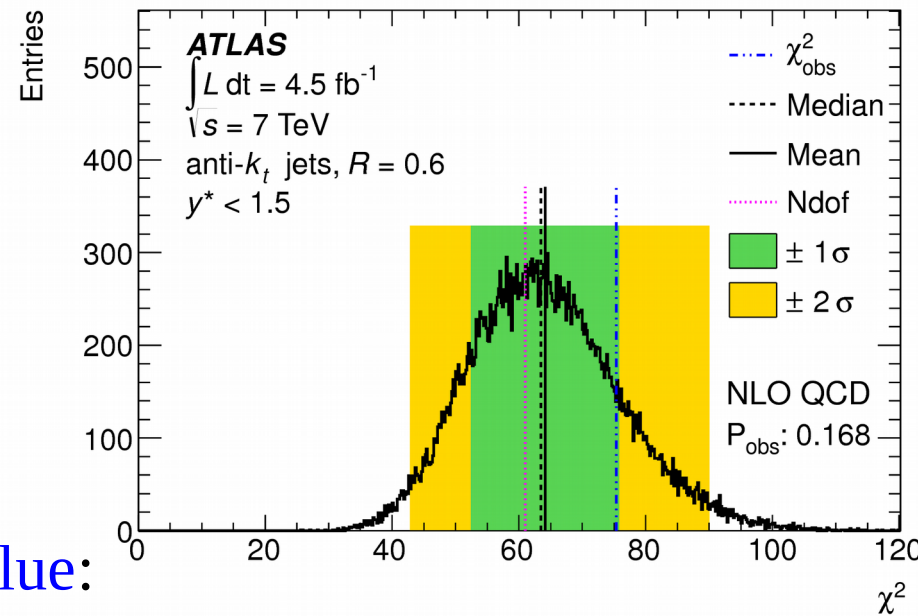
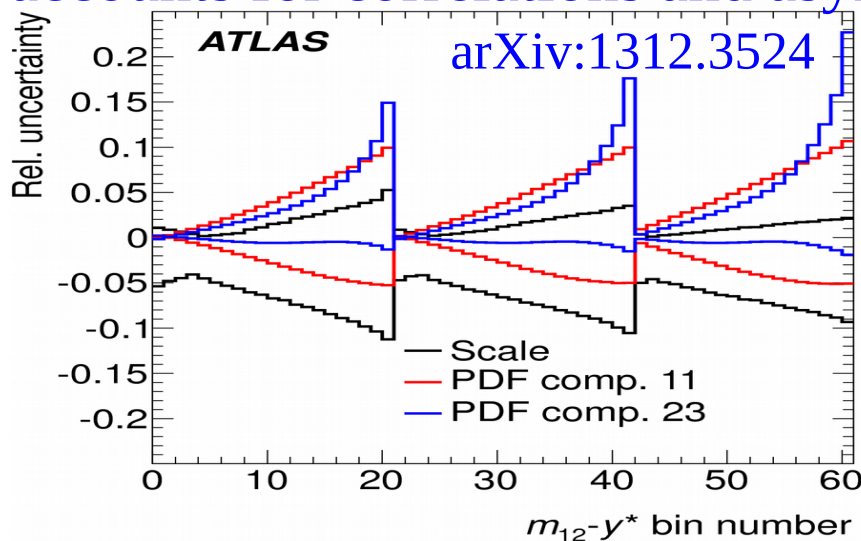
$$\chi^2(\mathbf{d}; \mathbf{t}) = \sum_i \left( \frac{d_i - t_i}{\sigma_i(t_i)} \right)^2 \longrightarrow \chi^2(\mathbf{d}; \mathbf{t}) = \sum_{i,j} (d_i - t_i) \cdot [C^{-1}(\mathbf{t})]_{ij} \cdot (d_j - t_j)$$



# Quantitative data/theory comparisons

- **Generalized**  $\chi^2(\mathbf{d}; \mathbf{t}) = \min_{\beta_a} \left\{ \sum_{i,j} \left[ d_i - \left( 1 + \sum_a \beta_a \cdot (\epsilon_a^\pm(\beta_a))_i \right) t_i \right] \cdot [C_{\text{su}}^{-1}(\mathbf{t})]_{ij} \cdot \left[ d_j - \left( 1 + \sum_a \beta_a \cdot (\epsilon_a^\pm(\beta_a))_j \right) t_j \right] + \sum_a \beta_a^2 \right\},$

→ accounts for correlations and asymmetries of uncertainties



- Using **frequentist method to compute p-value**:

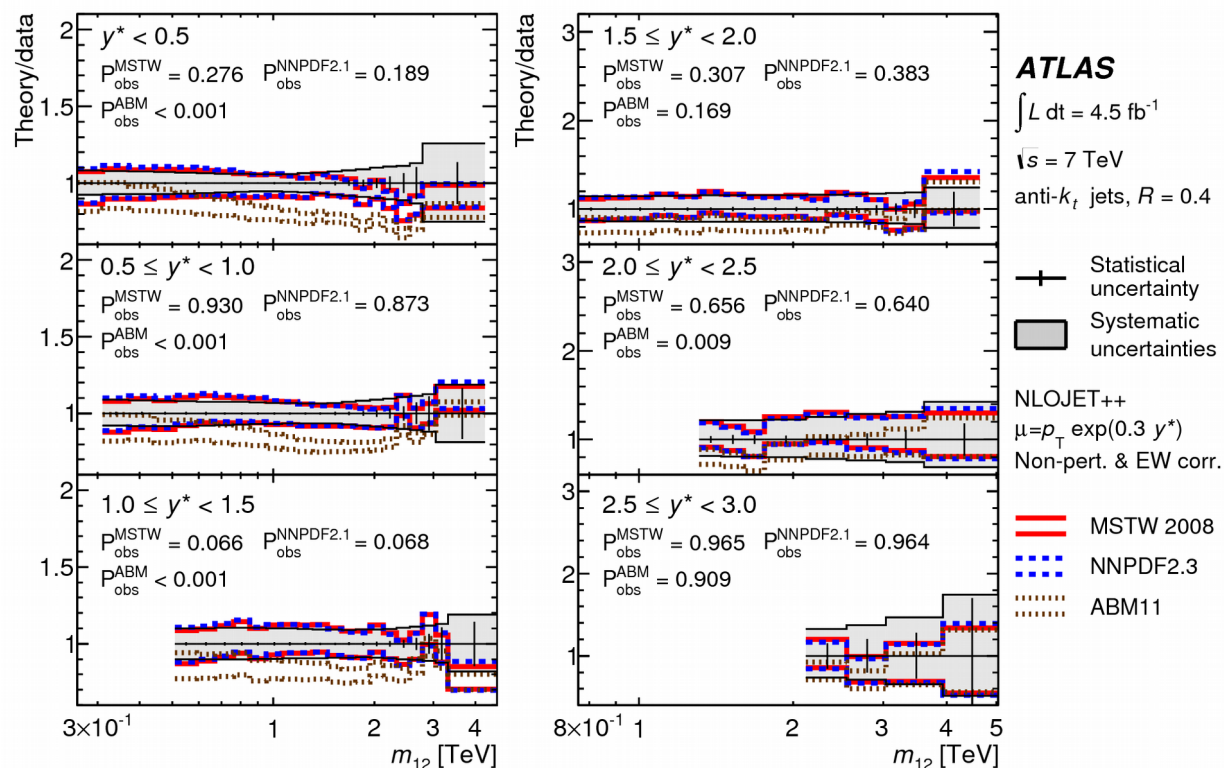
→ pseudo-experiments from theory prediction, with the full information on the uncertainties: **build the generalized  $\chi^2$  distribution (no assumption needed)**

→ observed  $\chi^2$  from the data/theory comparison

# PDF comparisons for dijets

arXiv:1312.3524

- Comparisons to MSTW2008, NNPDF2.3 and ABM11



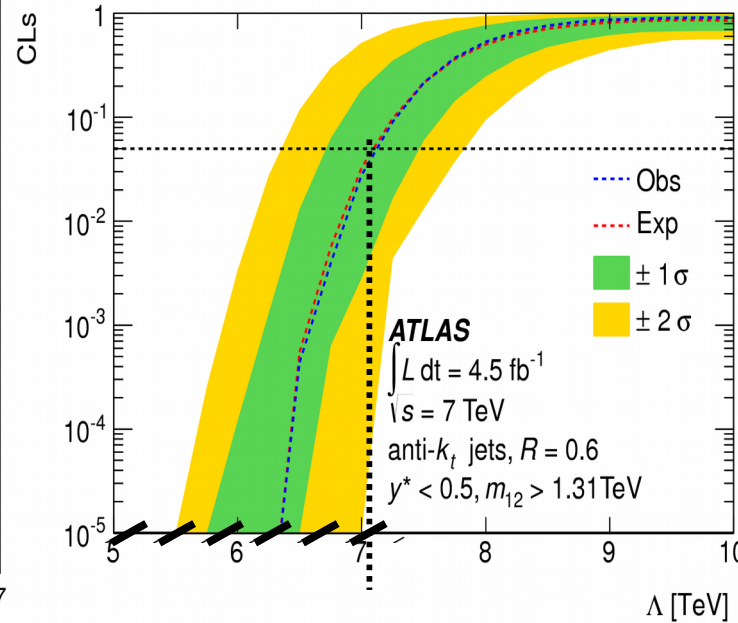
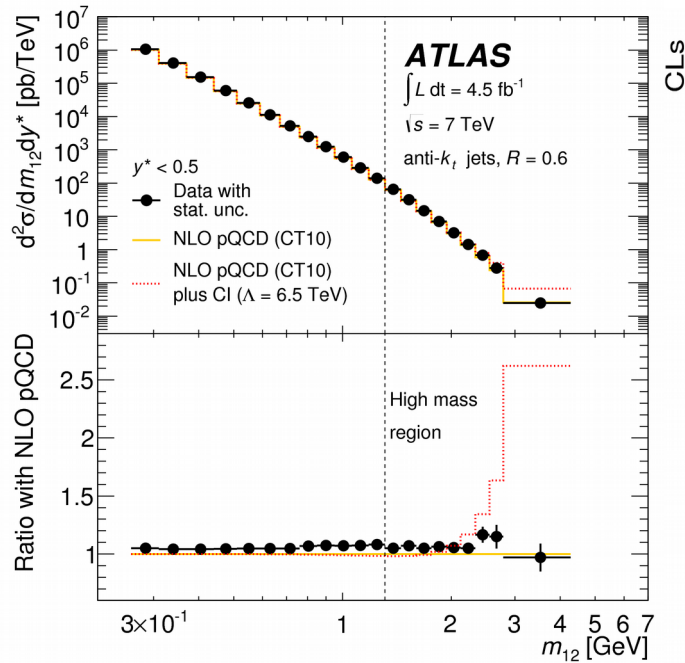
PDF set	$y^*$ ranges	mass range (full/high)	$P_{\text{obs}}$	
			$R = 0.4$	$R = 0.6$
CT10	$y^* < 0.5$	high	0.742	0.785
	$y^* < 1.5$	high	0.080	0.066
	$y^* < 1.5$	full	0.324	0.168
HERAPDF1.5	$y^* < 0.5$	high	0.688	0.504
	$y^* < 1.5$	high	0.025	0.007
	$y^* < 1.5$	full	0.137	0.025
MSTW 2008	$y^* < 0.5$	high	0.328	0.533
	$y^* < 1.5$	high	0.167	0.183
	$y^* < 1.5$	full	0.470	0.352
NNPDF2.1	$y^* < 0.5$	high	0.405	0.568
	$y^* < 1.5$	high	0.151	0.125
	$y^* < 1.5$	full	0.431	0.242
ABM11	$y^* < 0.5$	high	0.024	$< 10^{-3}$
	$y^* < 1.5$	high	$< 10^{-3}$	$< 10^{-3}$
	$y^* < 1.5$	full	$< 10^{-3}$	$< 10^{-3}$

- Sensitivity to PDFs:** level of agreement strongly depends on the PDF set and phase-space region

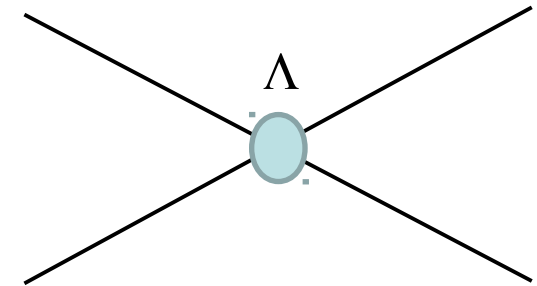
→ **Valuable experimental inputs to constrain proton PDFs:** Published information on **cross-sections & uncertainties, with their correlations and asymmetries**

# Limits on New Physics using unfolded distributions

- Explore BSM physics directly at particle level



Contact Interaction Model (CI)  
 New force mediated by heavy particle



arXiv:1312.3524

- Full frequentist analysis (CLs), with generalized  $\chi^2$  as test statistic  
 → accounts for correlations and asymmetries of uncertainties (stat. & syst.)
- Limits similar to the ones obtained by dedicated searches (comparing reco-level data with theory predictions folded with detector effects)

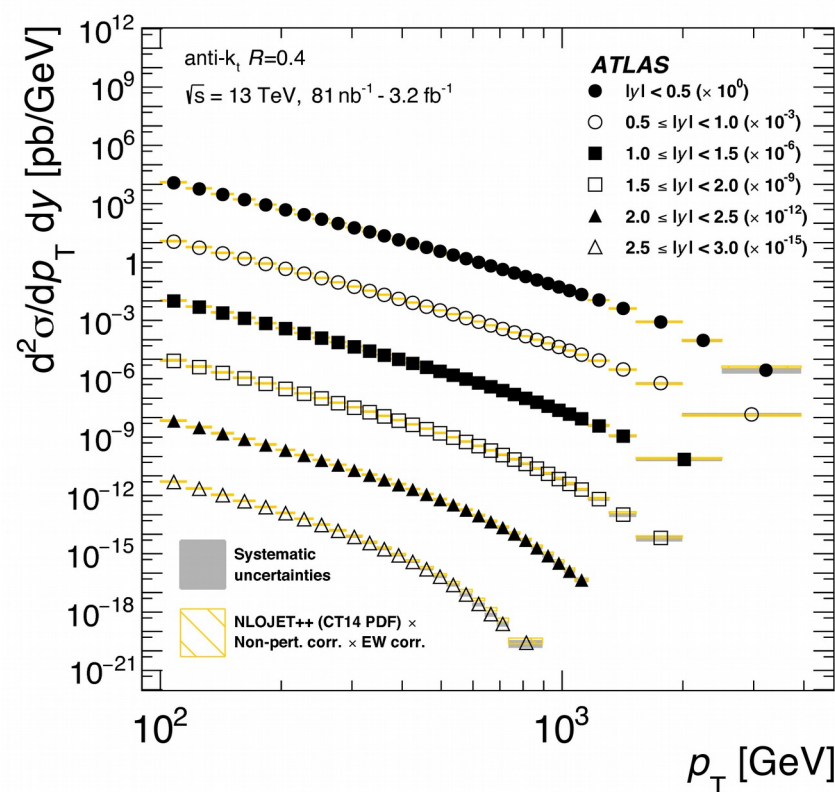
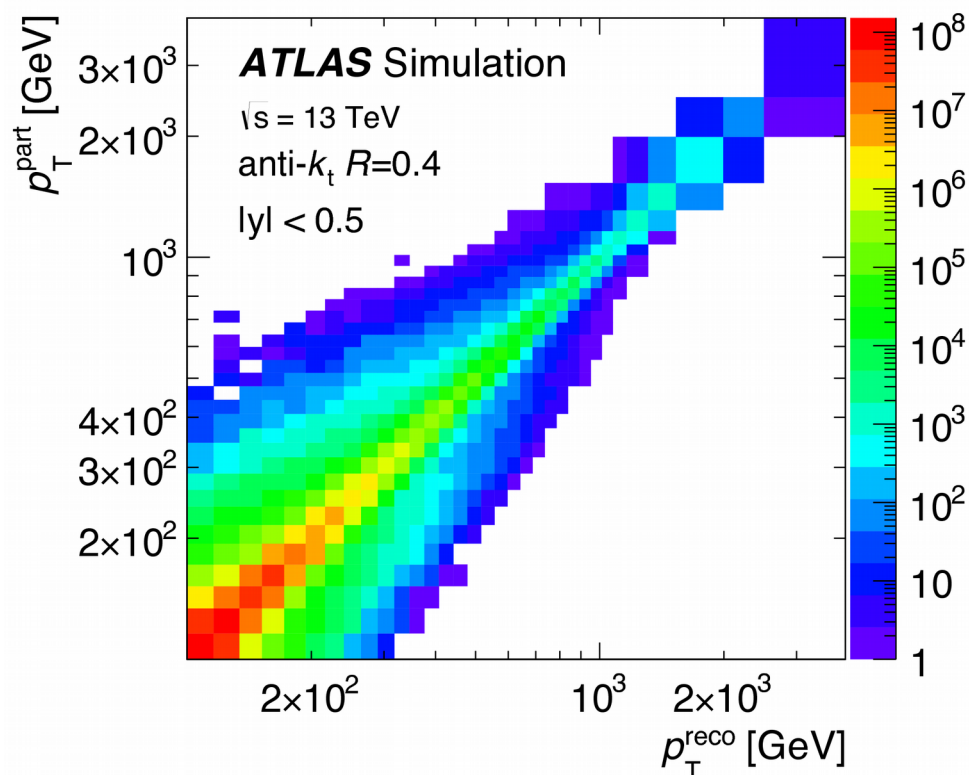
# nD Unfolding

# When to perform an unfolding in nD ?

- Decision to perform an unfolding in nD based on:
  - Observables of interest
  - Resolution for the observables of interest
  - Impact of merging of bins of one observable on the average detector response for another one ( see discussion on “hidden variables” )
  - Reduced statistics per bin when increasing the number of dimensions / reducing bin size:
    - having more bins enhances the amount of available information
    - some unfolding methods can be impacted in various ways by effects of low statistics

# nD unfolding – resolution effects mainly on 1 observable

- Several 2D studies (e.g. inclusive jets Xsec) where the resolution effects mainly impact one observable ( $p_T$ ), but have little impact on the other one ( $|y|$ )
  - 1D unfolding for several  $p_T$  slices
  - small migrations in  $|y|$  accounted for through efficiency corrections



# nD unfolding – resolution effects on several observables

- In general, even when resolution effects impact several observables, one can convert a nD unfolding exercise into 1D: bin re-ordering
- Care needed for adapting regularization methods
  - straight-forward for Bayesian-inspired iterative method, IDS, etc.  
(methods not directly sensitive to the relative position of the bins in the physical space; sensitive only to the amount of migrations between bins)
  - regularization method for SVD has to be adapted (average curvature directly sensitive to the relative position of the bins in the physical space; examples where the method has been adapted to 2D do exist)