Imperial College London

PHYSTAT-Systematics 2021 review: Physicist's view

Nicholas Wardle – Imperial College London

BIRS: Systematic Effects and Nuisance Parameters in Particle Physics Data Analyses

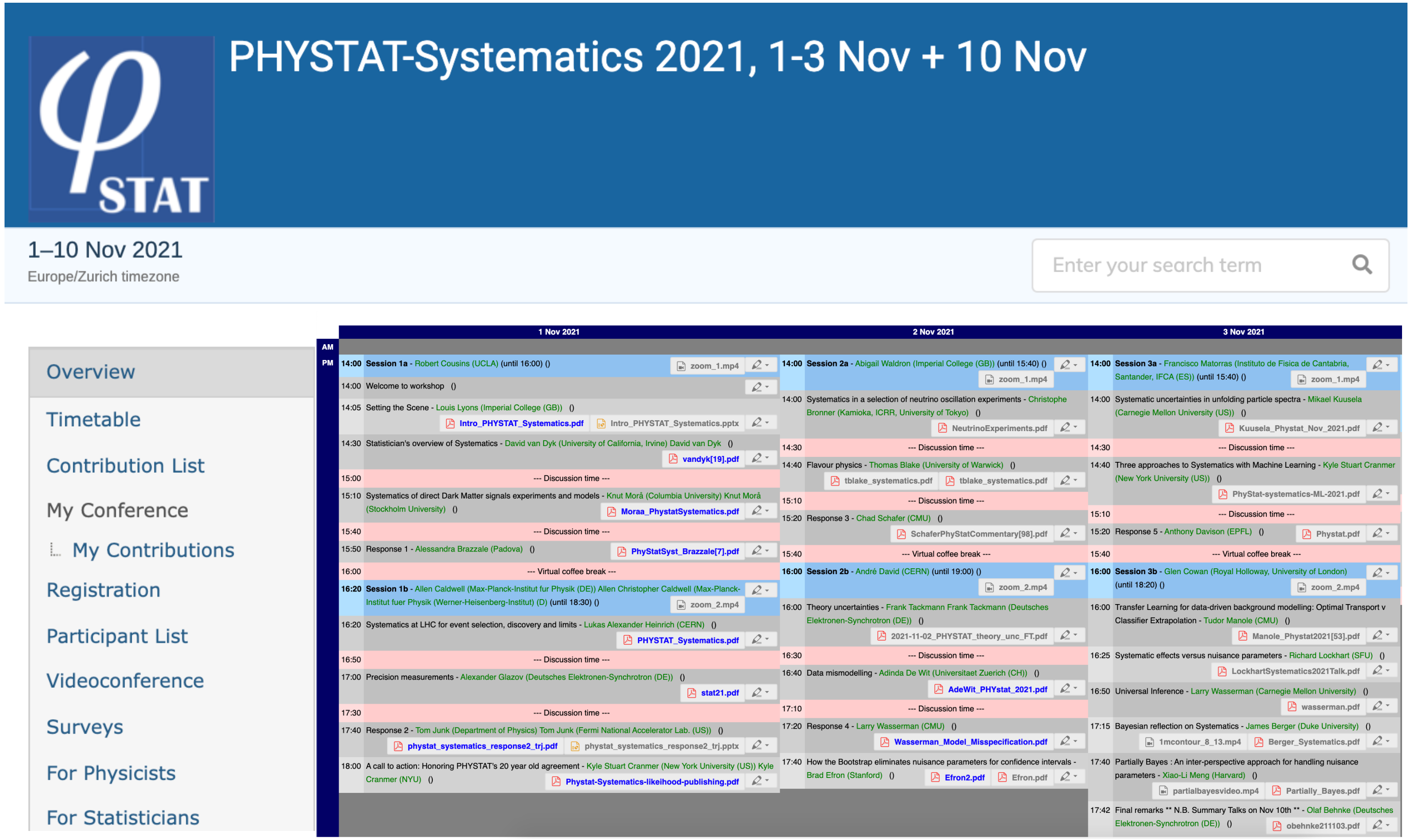
24/04/2023

PHYSTAT-Systematics Workshop

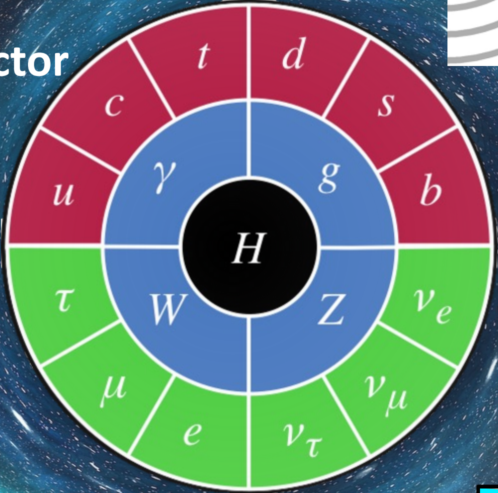
A remote Workshop devoted to the way systematic uncertainties are incorporated in data analyses in Particle Physics.

1st – 3rd November 2021: Three afternoons presentations/discussions/responses. Contributions from Physicists & Statisticians

All contributions & session recordings available at : https://indico.cern.ch/event/1051224/



(some of) The experimental particle physics landscape



Dark sector

Astroparticle/cosmology

Neutrino physics

Energy frontier (my area)

Flavor physics

In this talk I'll try to cover some common issues ...

The Good, The Bad and The Ugly*

- **The Good:** Your own calibrations → basically statistical
- **The Bad:** Using other peoples results, poorly modelled data or analysis technique, model assumptions, …
- **The Ugly**: Different theoretical estimates, theory with limited number of terms, …

**Pekka Sinervo – PHYSTAT 2003*

Sources of systematic uncertainty

Huge array of sources of systematic uncertainty in particle physics experiments (see individual talks for details)…

- Dark matter experiments : **Systematics of direct Dark Matter signals experiments and models – Knut Moraa**
- Searches at the LHC : **Systematics at LHC for event selection, discovery and limits – Lukas Heinrich**
- Precision measurements at colliders & beam expts. : **Precision measurements – Alexander Glazov**
- Neutrino oscillation experiments : **Systematics in a selection of neutrino oscillation experiments – Christophe Bronner**
- Searches and measurements in flavour physics **: Flavour Physics – Thomas Blake**
- Uncertainties in Theory calculations **: Theory uncertainties – Frank Tackmann**

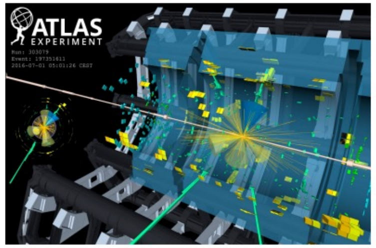
Sources of systematic uncertainty

Huge array of sources of systematic uncertainty in particle physics experiments (see individual talks for details)…

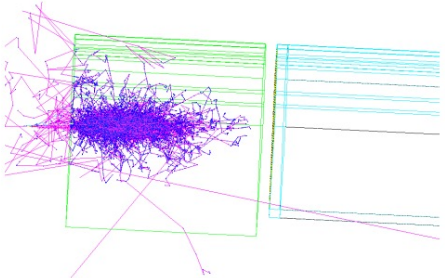
- Dark matter experiments : Systematics of direct Dark Matter signals experiments and models **– Knut Moraa**
- Searches at the LHC : Systematics at LHC for event selection, discovery and limits **– Lukas Heinrich**
- Precision measurements at colliders & beam expts. : Precision measurements **– Alexander Glazov**
- Neutrino oscillation experiments : Systematics in a selection of neutrino oscillation experiments **– Christophe Bronner**
- Searches and measurements in flavour physics **:** Flavour Physics **– Thomas Blake**
- Uncertainties in Theory calculations **:** Theory uncertainties **– Frank Tackmann**

$$p(\mathrm{data}\,|\,\theta) = \iiint \mathrm{d}z_d \mathrm{d}z_h \mathrm{d}z_p \; p(\mathrm{data}\,|\,z_p) \;\; p(z_d|z_h)\, p(z_h|z_p)\, p(z_p|\theta)$$

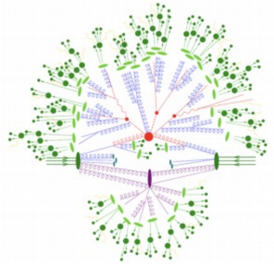
Lukas Heinrich



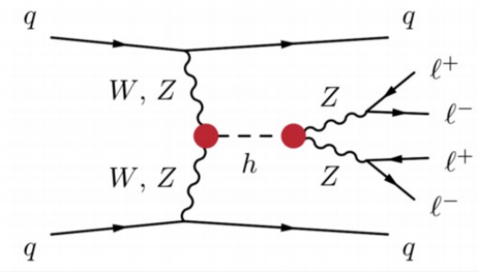
detector readout



detector interaction



stochastic evolution



hard process

Systematic Recipes

Typically (broadly) two strategies for including systematic uncertainties in limits, measurements, …

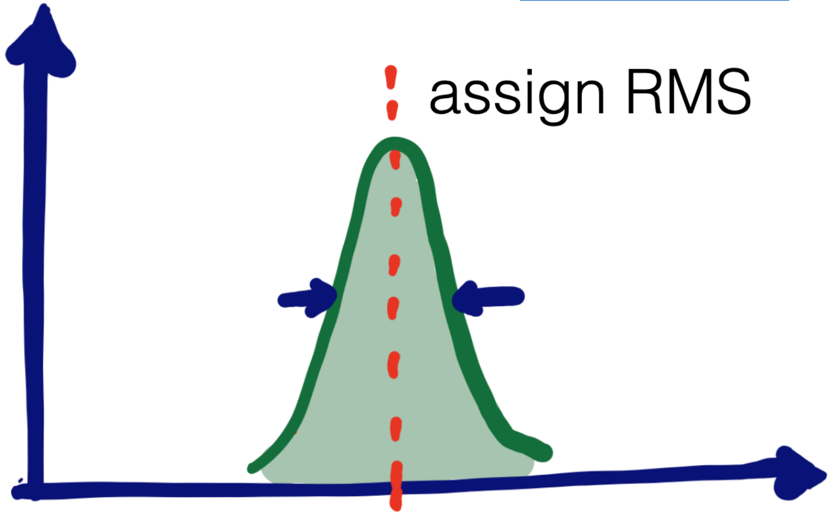
- Error propagation

→ Change a single parameter of the model / swap out the nominal model for an alternate one representing a systematic shift from particular source: One Parameter at a Time (OPAT)

→ Throw random toys representing systematic variations to determine spread of results → MC method

→ Difference in result(s) (measurement) quoted as systematic uncertainty due to that source(s)

Thomas Blake



Systematic Recipes

Typically (broadly) two strategies for including systematic uncertainties in limits, measurements, …

- Error propagation

→ Change a single parameter of the model / swap out the nominal model for an alternate one representing a systematic shift from particular source: One Parameter at a Time (OPAT)

→ Throw random toys representing systematic variations to determine spread of results → MC method

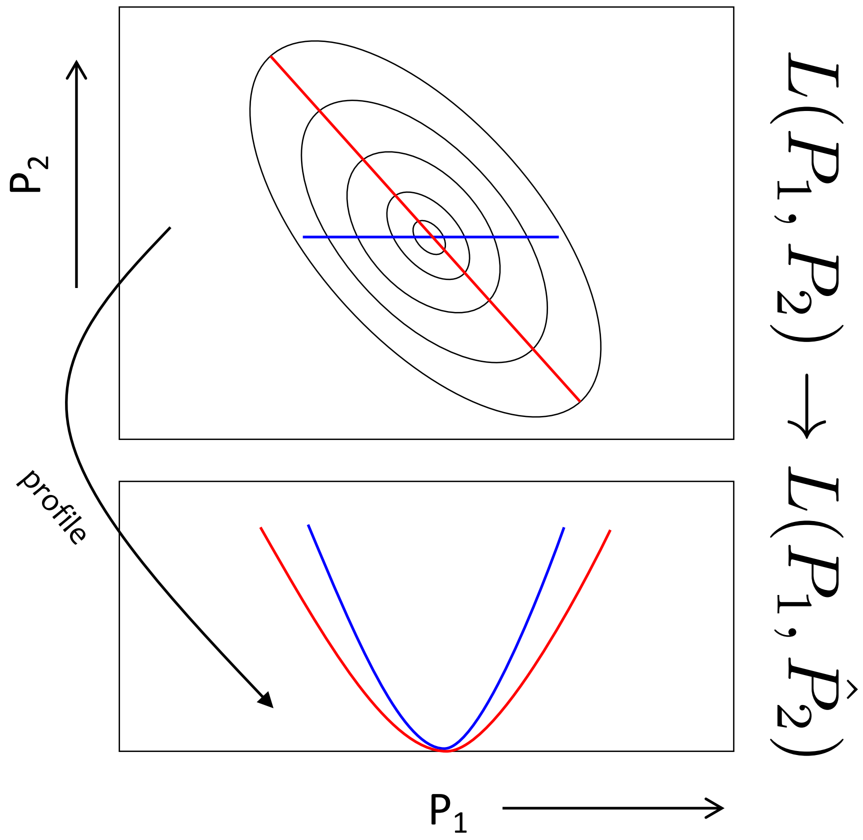
→ Difference in result(s) (measurement) quoted as systematic uncertainty due to that source(s)

- Likelihood based approach

→ Systematic uncertainties encoded as ***Nuisance Parameters*** in probability density model

→ Constraints / Priors for nuisance parameters often derived from one or more "*Preliminary analysis*"

→ Construct likelihood for "*Primary analysis*" and profile/marginalize over nuisance parameters



I leave the proper discussion of these approaches to Sara Algeri's (Statistician's view) talk!

Error propagation

OPAT : One Parameter At a Time

Uncertainty due to source "s_j" on measurement "μ_i" determined as

Alexander Glazov

$$\Gamma_{ij} = \mu_i(s_j^0 + \delta s_j) - \mu_i(s_j^0)$$

For independent sources, sum in quadrature (over i) to get total uncertainty on each measurement.

Good for simplicity of implementation but need to be careful about keeping track of signs

Error propagation

OPAT : One Parameter At a Time

Uncertainty due to source "s_j" on measurement "μ_i" determined as

Alexander Glazov

$$\Gamma_{ij} = \mu_i(s_j^0 + \delta s_j) - \mu_i(s_j^0)$$

For independent sources, sum in quadrature (over i) to get total uncertainty on each measurement.

Good for simplicity of implementation but need to be careful about keeping track of signs

MC method

Prepare random samples of systematic shifts according to some distribution (e.g $r \sim N(0,1)$)

$$s'_j = s_j^0 + r_j \delta s_j$$

Repeat measurements for each sample and estimate uncertainty according to

$$\sigma\mu_i = \sqrt{\frac{1}{N_r - 1}\sum_{k=1}^{N_r}(\mu_i^r - \overline{\mu_i^k})^2}\,.$$

Discussion points:

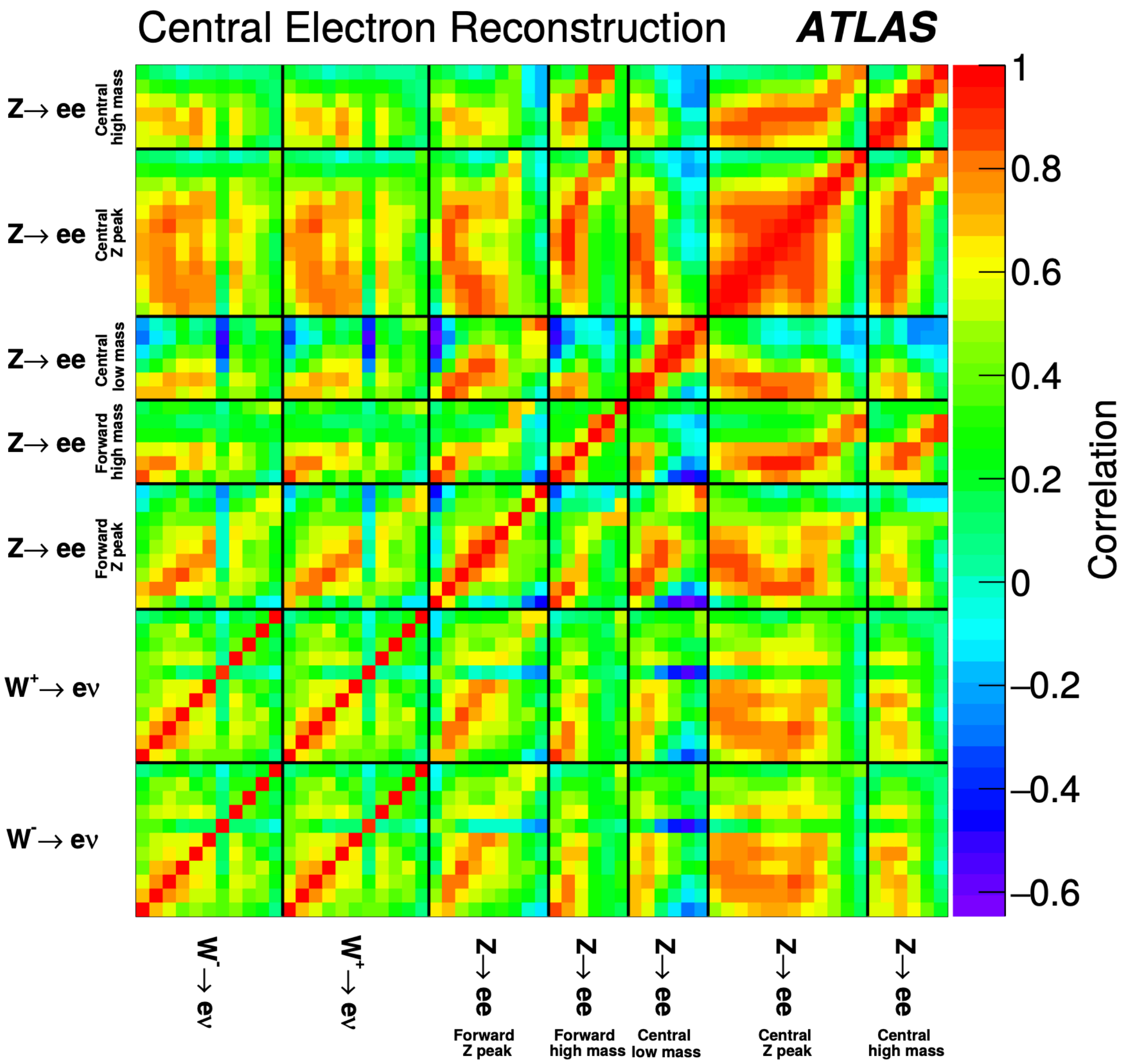
- *"Is providing covariance enough?"*

ATLAS differential X-sections

Alexander Glazov

MC variations also used to determine correlation between reconstruction efficiency measurements

→ Differential Z&W boson cross-sections rely on detailed understanding of correlations across different rapidity regions



ATLAS, EPJC 77 (2017) 367

Nuisance parameters

Nuisance parameters **ν** "built into" statistical model (probability density)

$$p(\mathrm{data}|\theta) \to p(\mathrm{data}|\theta, \nu)$$

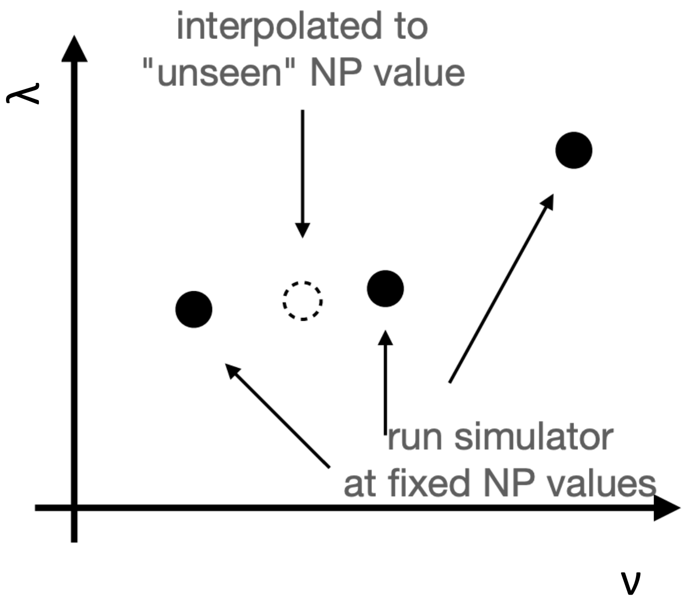
→ Need to parameterize effects of nuisance parameters on density

E.g let

$$p = \frac{\lambda^n e^{-\lambda}}{n!}$$

Build nuisance parameter effects from shifting source and calculating size of effect on λ

$$\lambda(\nu) = \lambda_0(1+k)^\nu$$



Nuisance parameters

Nuisance parameters **ν** "built into" statistical model (probability density)

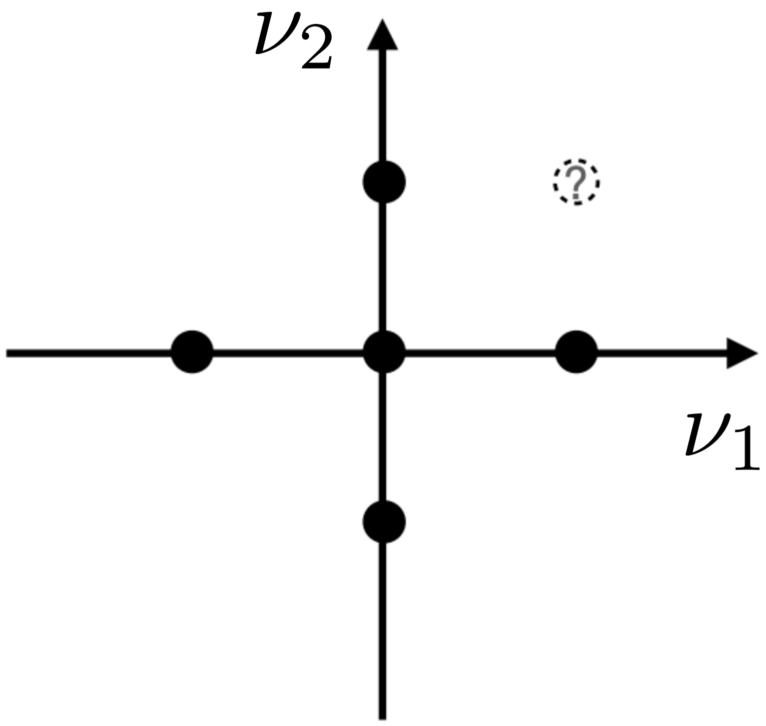
$$p(\mathrm{data}|\theta) \rightarrow p(\mathrm{data}|\theta, \nu)$$

→ Need to parameterize effects of nuisance parameters on density

E.g let $$p = \frac{\lambda^n e^{-\lambda}}{n!}$$

Suppose we want to model the effect of multiple nuisance parameters?
Typical choice is to factorize effects as multiplicative terms

$$\lambda(\nu) = \lambda_0 (1 + k_1)^{\nu_1} (1 + k_2)^{\nu_2}$$

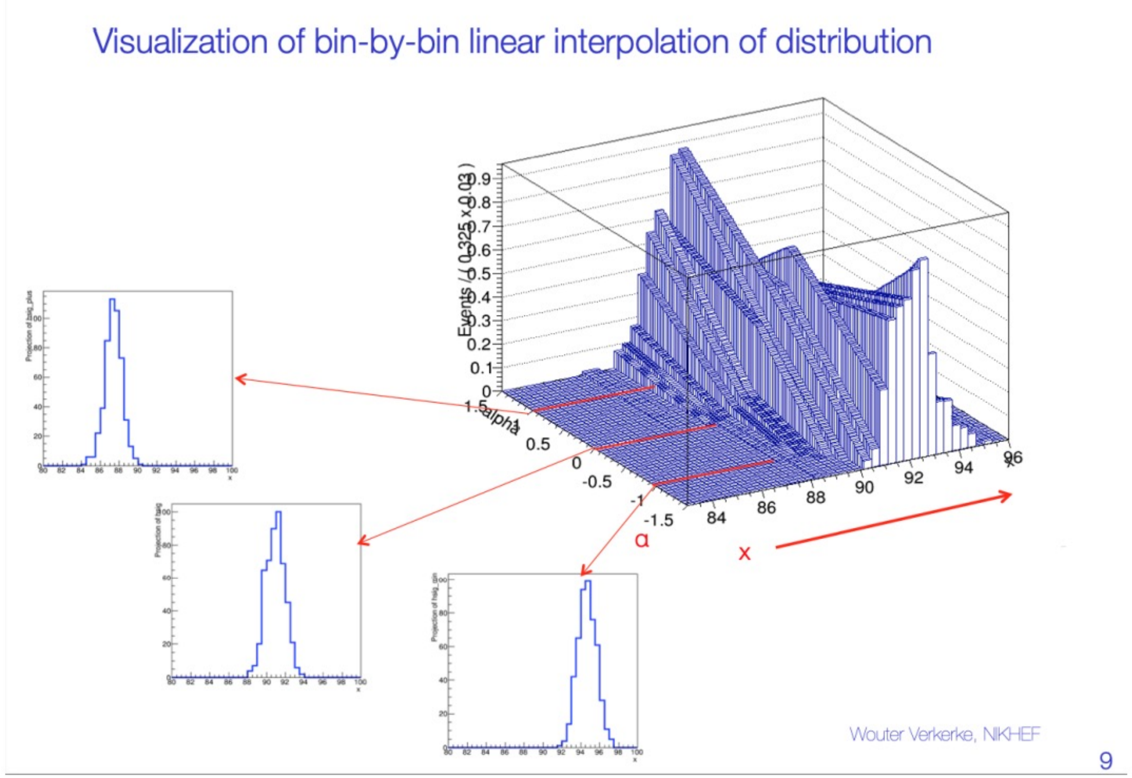


Discussion points:
- "How good is this factorisation assumption?"
- "How can we effectively test it when we have O(100) nuisance parameters?"

Lukas Heinrich

Template morphing

Adinda De Wit



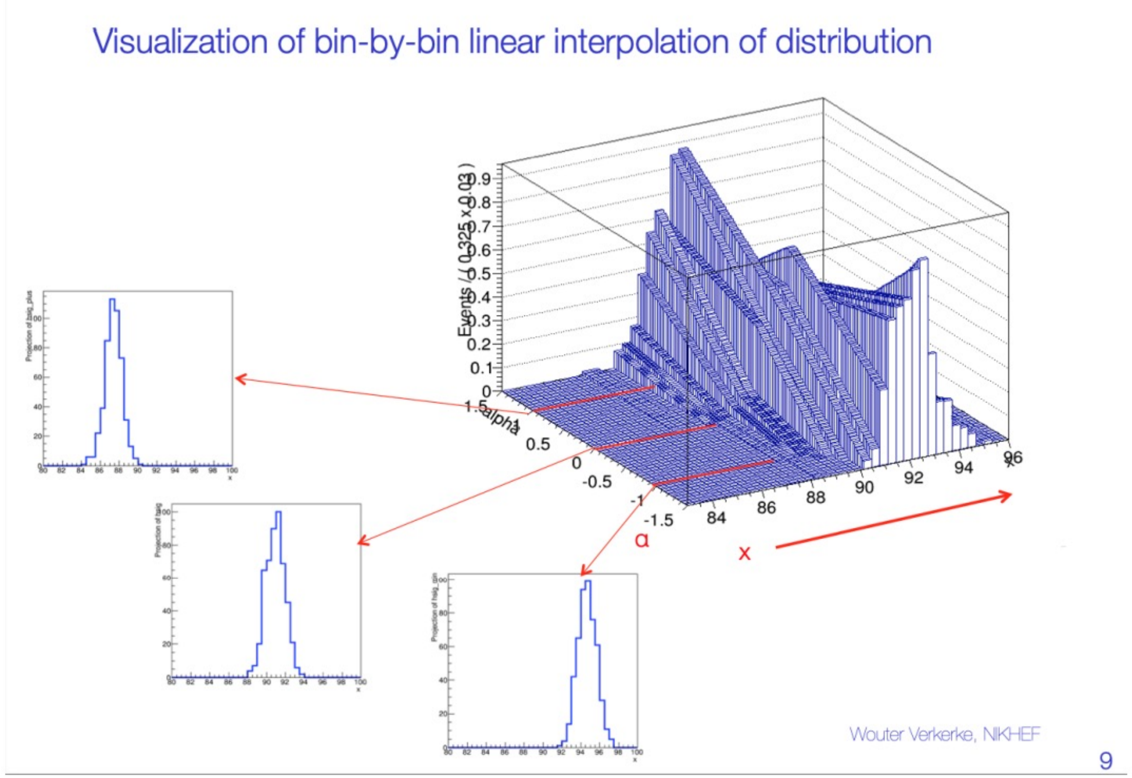
Linear interpolation* between p(x|α) at fixed values of α yields empirical parameterisation.

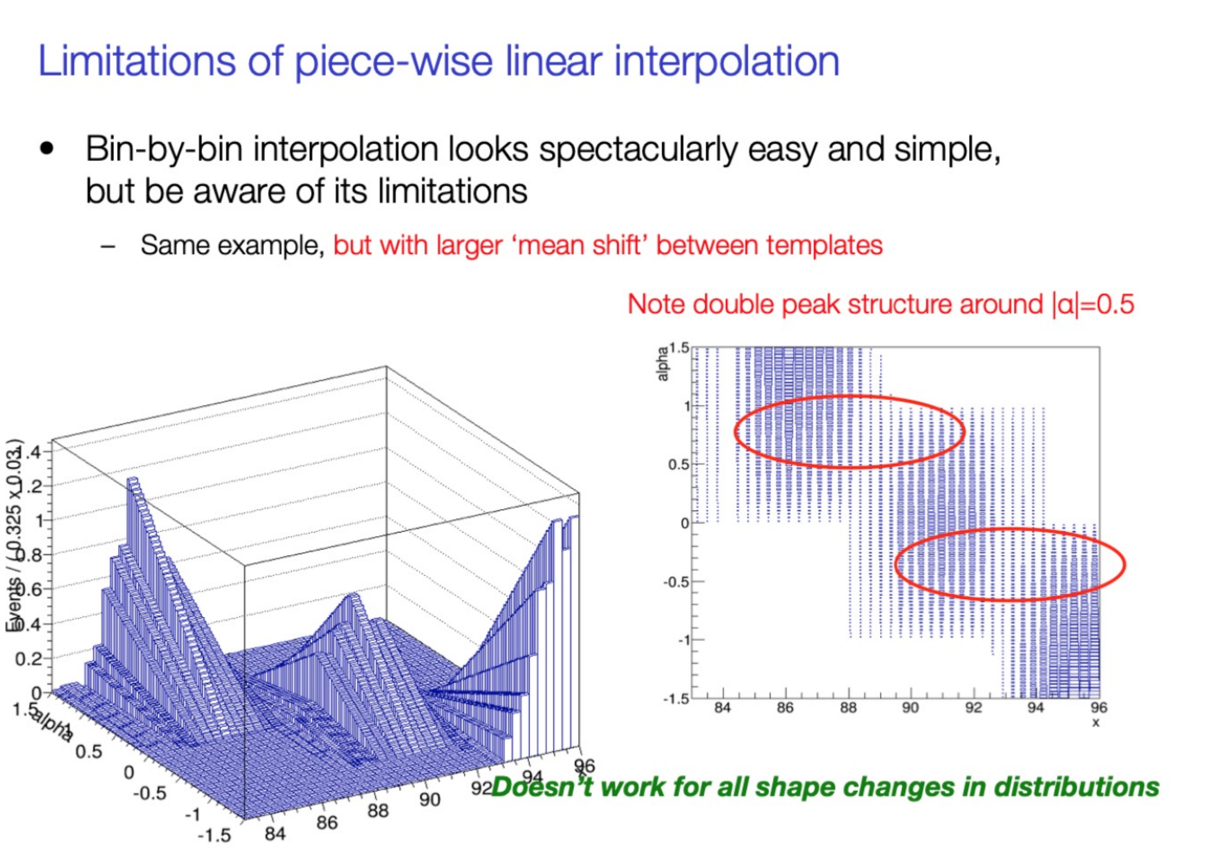
* Note, much more common to use polynomial interpolation (see backup)

Template morphing

[Adinda De Wit](#)

Linear interpolation* between p(x|α) at fixed values of α yields empirical parameterisation.





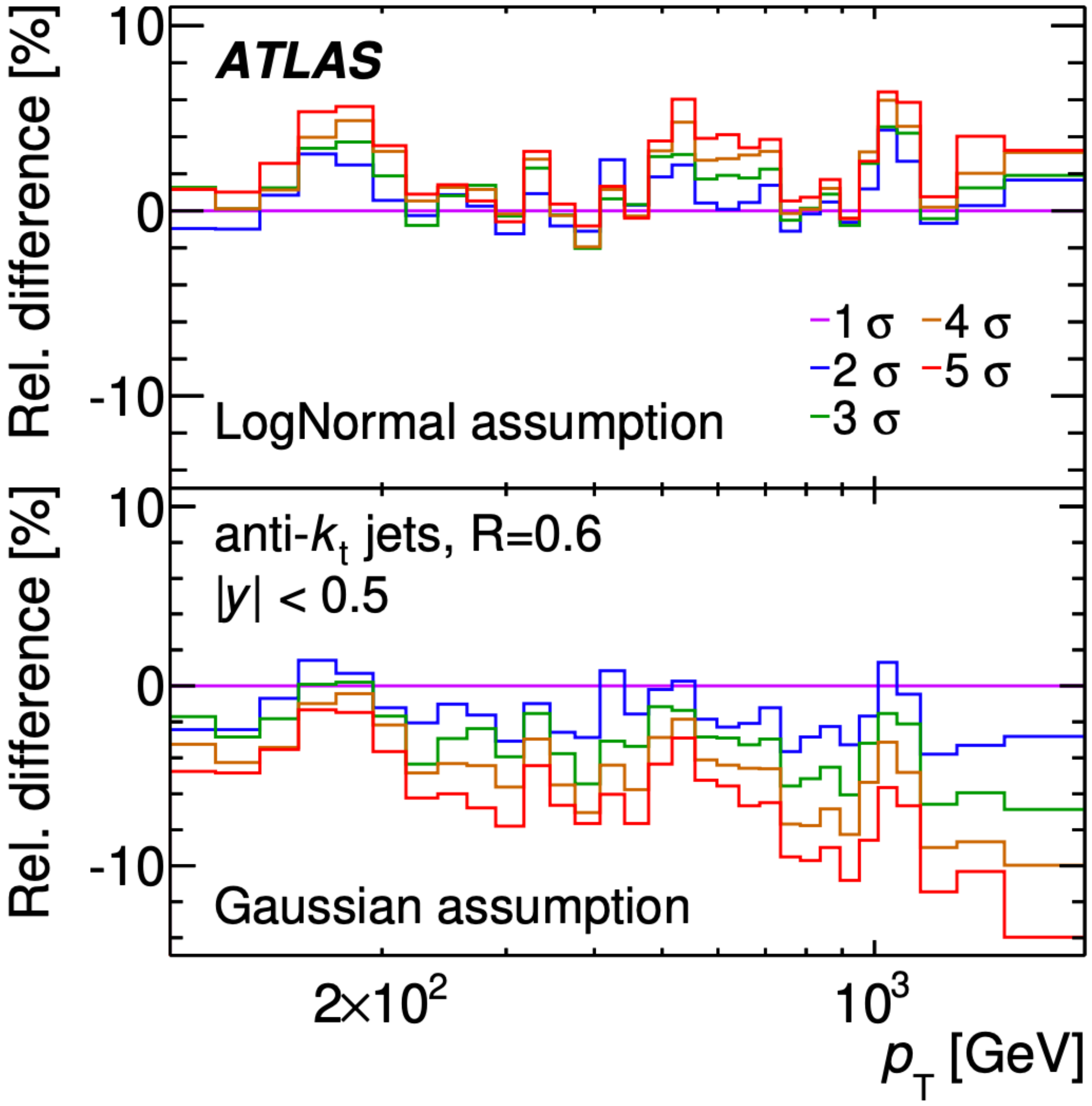
Cannot always rely on this approach (vertical interpolation) → e.g very typical in large shifts of the mean

* Note, much more common to use polynomial interpolation (see backup)

Non-Gaussian effects

Precision measurements need precision modeling

- Often use fixed points (-1σ, 0σ, +1σ) for interpolation
- Assume Gaussian behavior 2σ = 2 x 1σ
- Not necessarily true that this holds for all kinds of uncertainties
    - E.g $\lambda(\nu) = \lambda_0(1+\delta\nu)$ vs $\lambda(\nu) = \lambda_0(1+\delta)^\nu$



[Alexander Glazov](#) JHEP 02 (2015) 153

Non-Gaussian effects

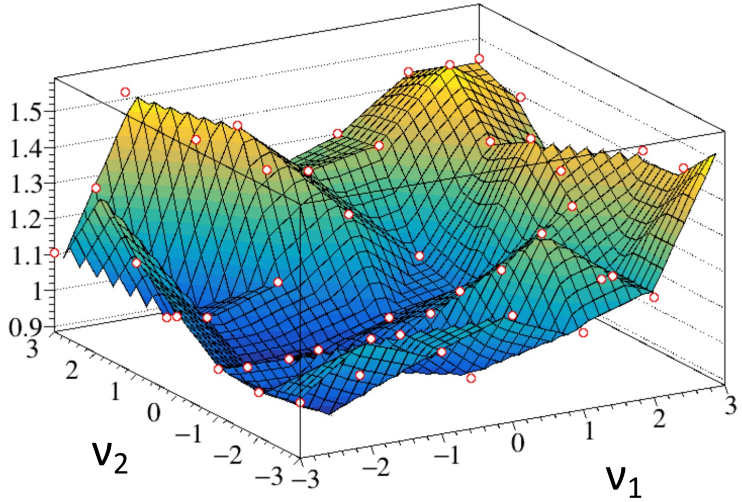
Precision measurements need precision modeling

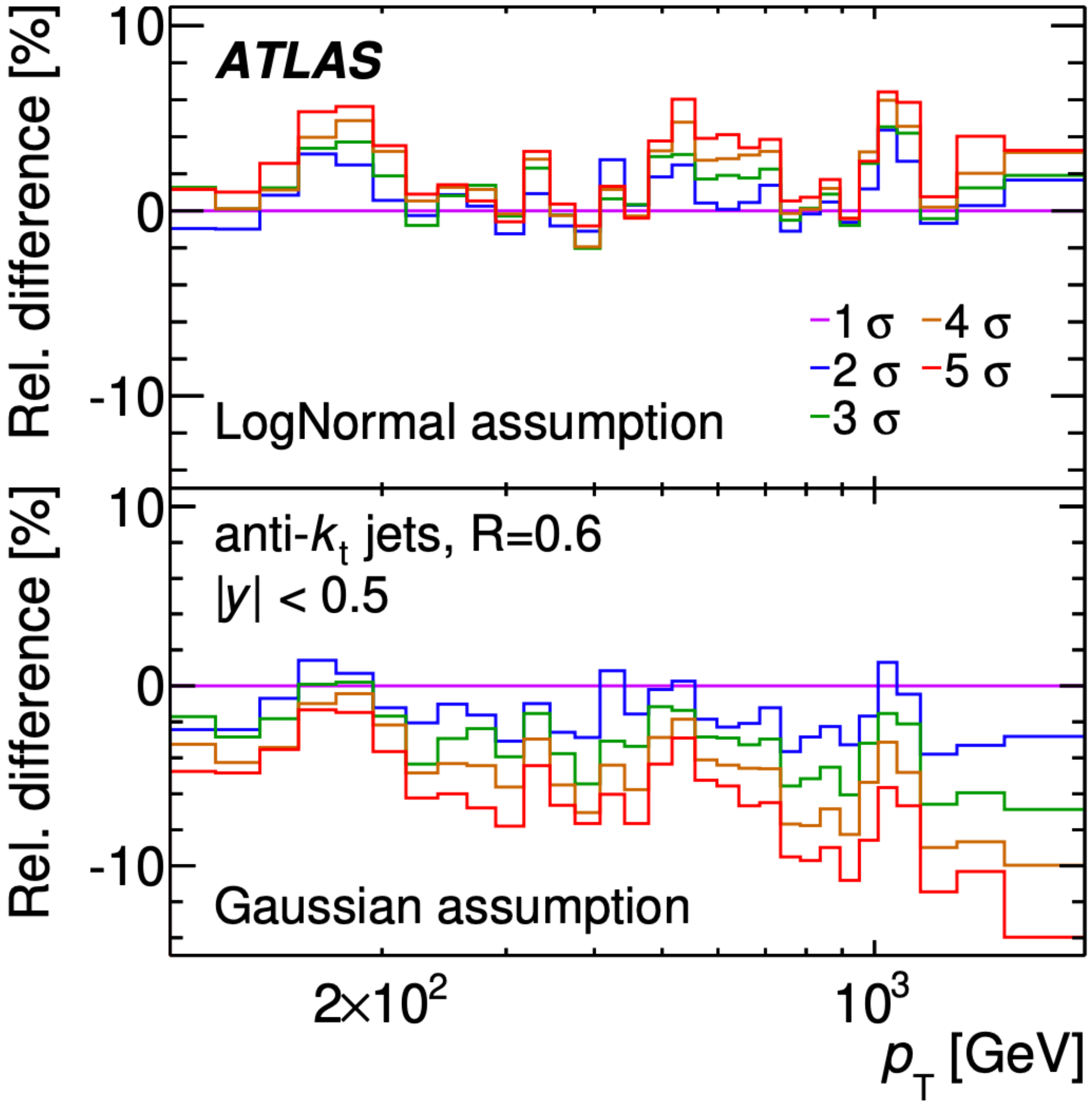
- Often use fixed points (-1σ, 0σ, +1σ) for interpolation
- Assume Gaussian behavior 2σ = 2 x 1σ
- Not necessarily true that this holds for all kinds of uncertainties
  - E.g $\lambda(\nu) = \lambda_0(1+\delta\nu)$ vs $\lambda(\nu) = \lambda_0(1+\delta)^{\nu}$

Discussion points:

- *"Should we sample many different parameter values to build suitable parameterisation?"*
- *"Are there smarter ways (eg GPs/ML?) to automate?"*

Christophe Bronner



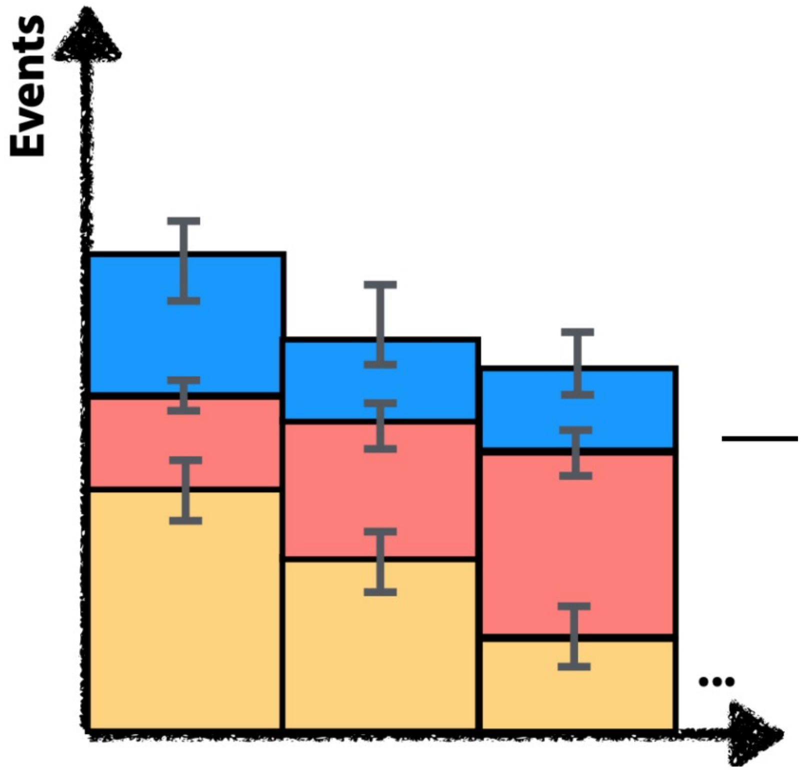


Alexander Glazov

JHEP 02 (2015) 153

Monte Carlo statistics uncertainties

Adinda De Wit



Barlow-Beeston method
Comput. Phys. Comun 77 (1993) 219

Often rely heavily on Monte Carlo event simulation to estimate probability densities and construct likelihoods

Generating MC can be CPU expensive so need to account for limited MC sample size when performing statistical analysis

Statistical in nature → in principle easy to model (need to account for weights in MC which are not always equal for a given sample)

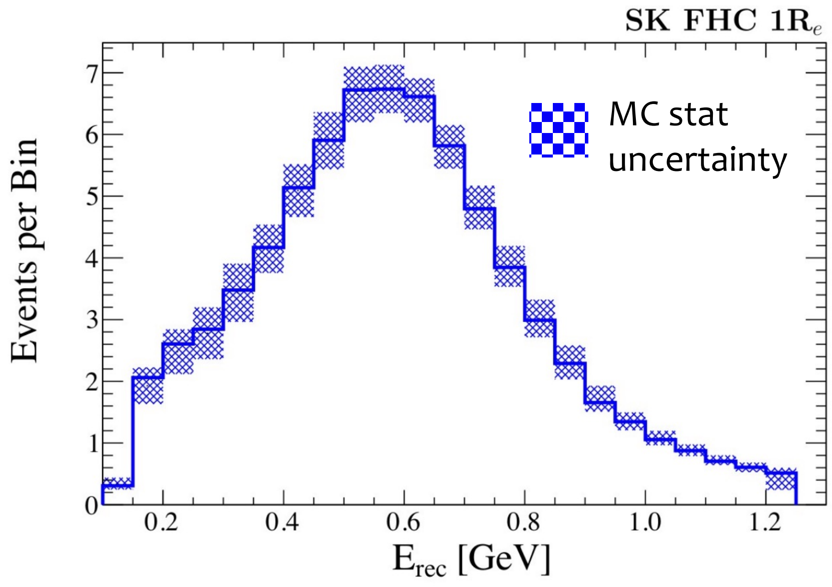
Tricks such as the Barlow-Beeston method help reduce impact on statistical analysis CPU time due to flood of additional nuisance parameters

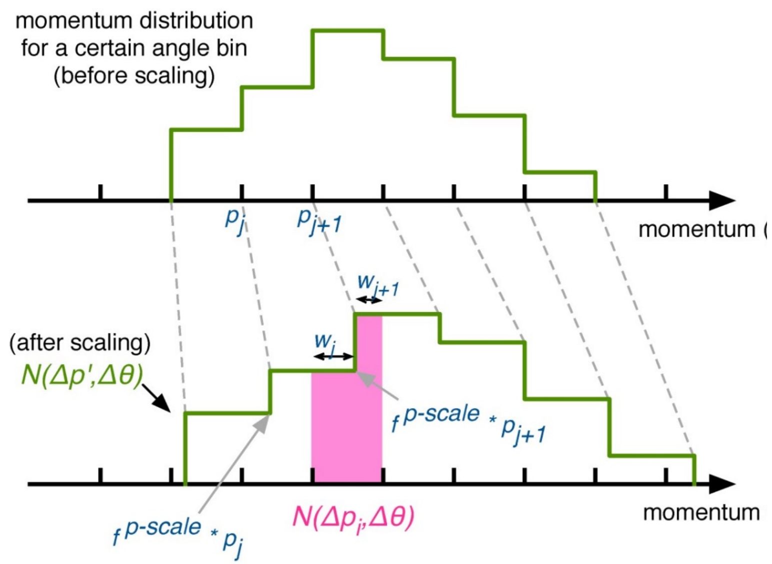
Simulation statistics

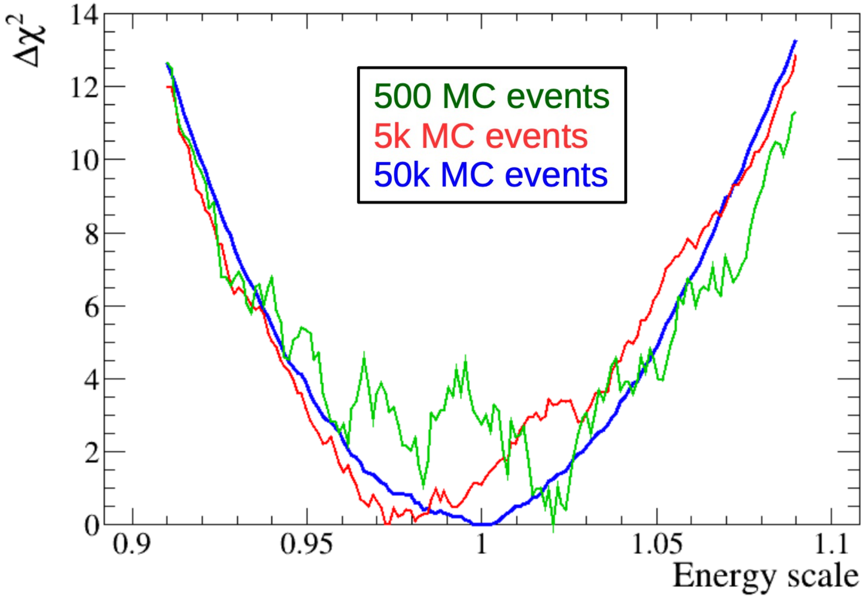
Nominal distributions (probability density) often determined using Monte Carlo Simulated events
→ Also use simulation directly to determine variation in different bins of some observable

e.g energy scale for momentum distribution in T2K

Christophe Bronner







Shift energy scale in simulation and calculate migration between neighboring bins

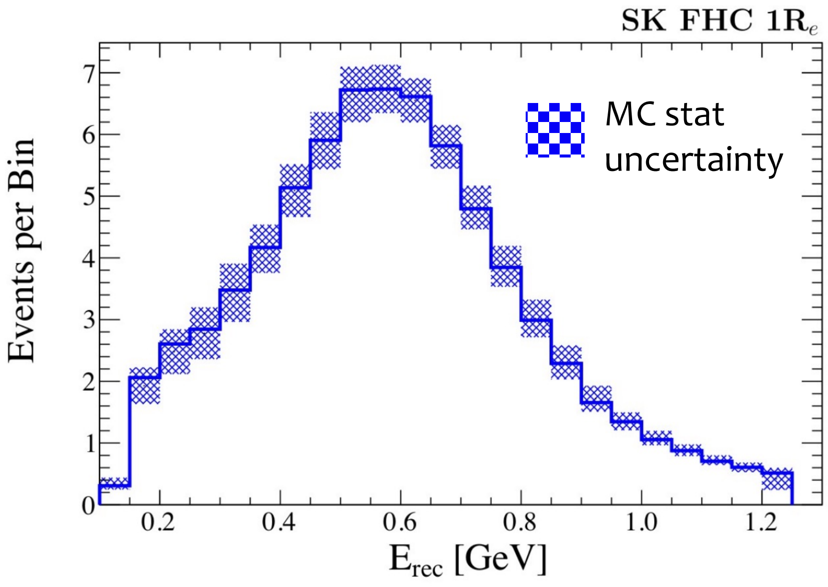
Jumps in χ^2 due to events jumping between bins

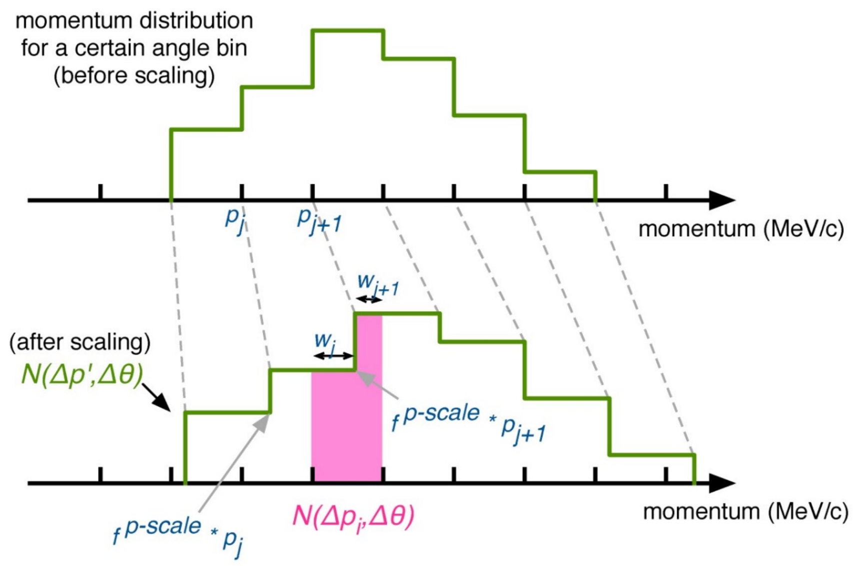
Simulation statistics

Nominal distributions (probability density) often determined using Monte Carlo Simulated events
→ Also use simulation directly to determine variation in different bins of some observable

e.g energy scale for momentum distribution in T2K

[Christophe Bronner]



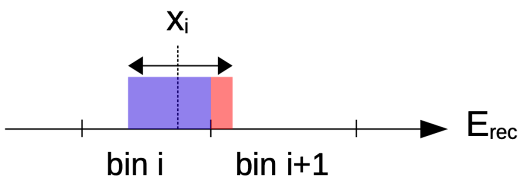


Shift energy scale in simulation and calculate migration between neighboring bins

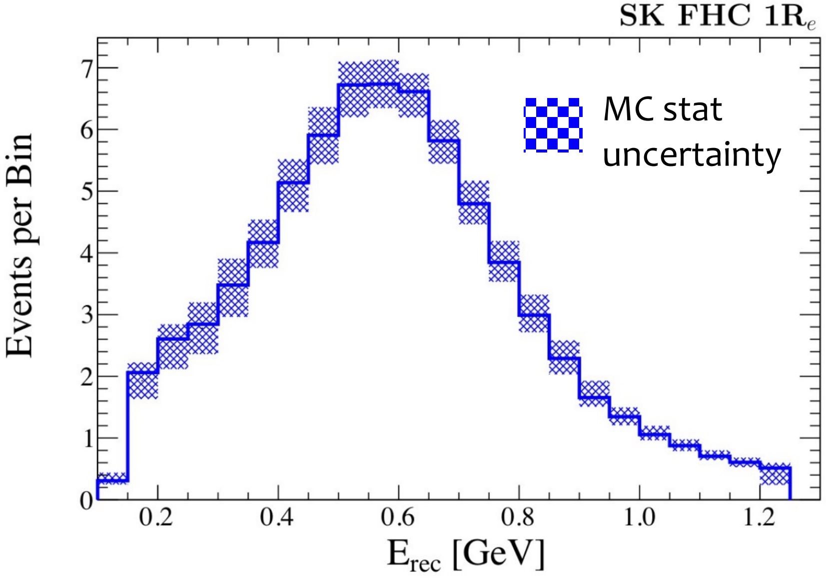
Simulation statistics

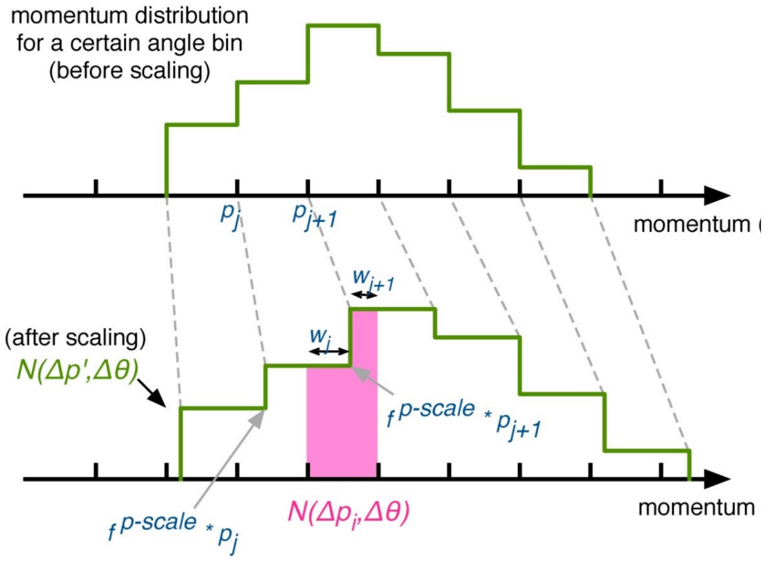
Nominal distributions (probability density) often determined using Monte Carlo Simulated events
→ Also use simulation directly to determine variation in different bins of some observable

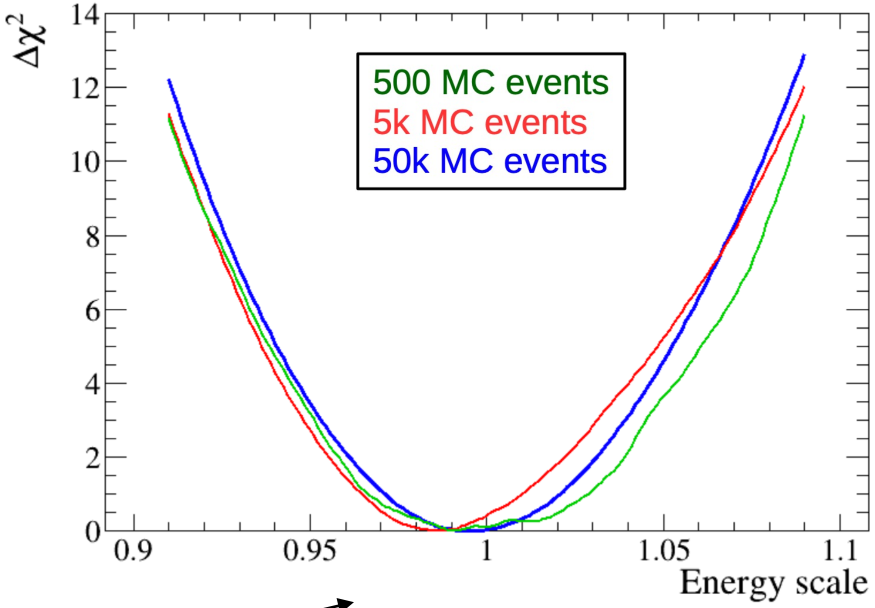
e.g energy scale for momentum distribution in T2K



Christophe Bronner







Shift energy scale in simulation and calculate migration between neighboring bins

Regularisation can help smooth out these effects: e.g distribute events across-bins (assign "width" to each event)

Discrete choices

In some cases, not obvious how to construct parameterization of a systematic uncertainty - Two point systematic
→ Typically related to a model choice : eg Pythia vs Herwig for Parton shower model

How do we account for this uncertainty ?

→ Bayesian might assign equal prior to each and marginalize, Frequentist might construct "average model"

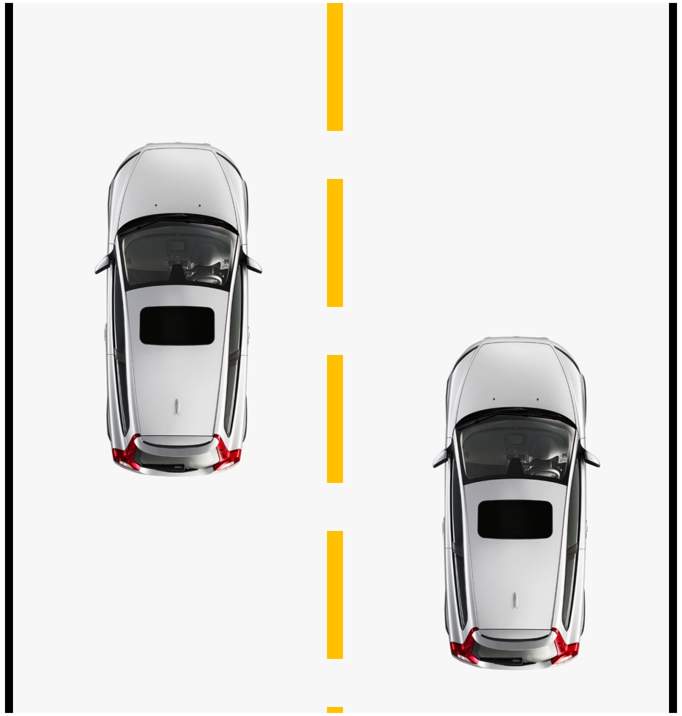
Discrete choices

In some cases, not obvious how to construct parameterization of a systematic uncertainty - Two point systematic
→ Typically related to a model choice : eg Pythia vs Herwig for Parton shower model

How do we account for this uncertainty ?

→ Bayesian might assign equal prior to each and marginalize, Frequentist might construct "average model"

→ Average model might not necessarily correspond to something meaningful



Phillip Litchfield two lane traffic example

Discrete choices

In some cases, not obvious how to construct parameterization of a systematic uncertainty - Two point systematic
→ Typically related to a model choice : eg Pythia vs Herwig for Parton shower model

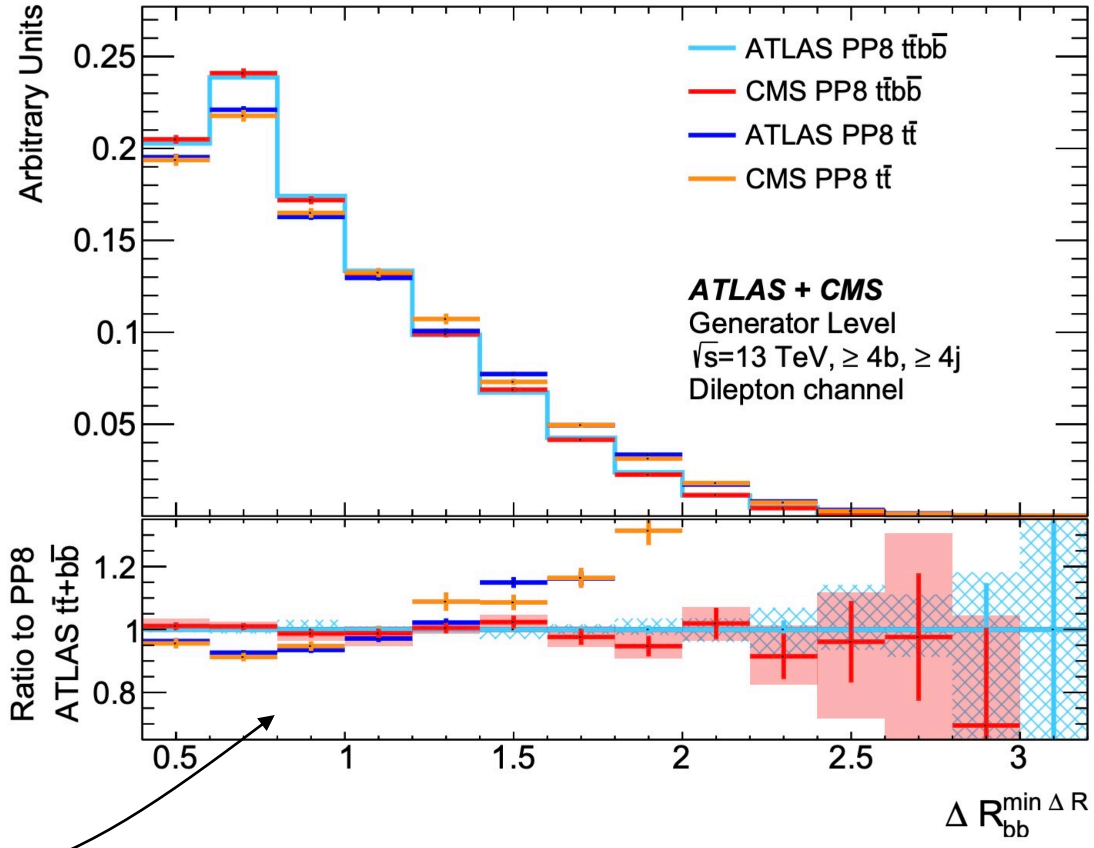
How do we account for this uncertainty ?

→ Bayesian might assign equal prior to each and marginalize, Frequentist might construct "average model"

→ Average model might not necessarily correspond to something meaningful

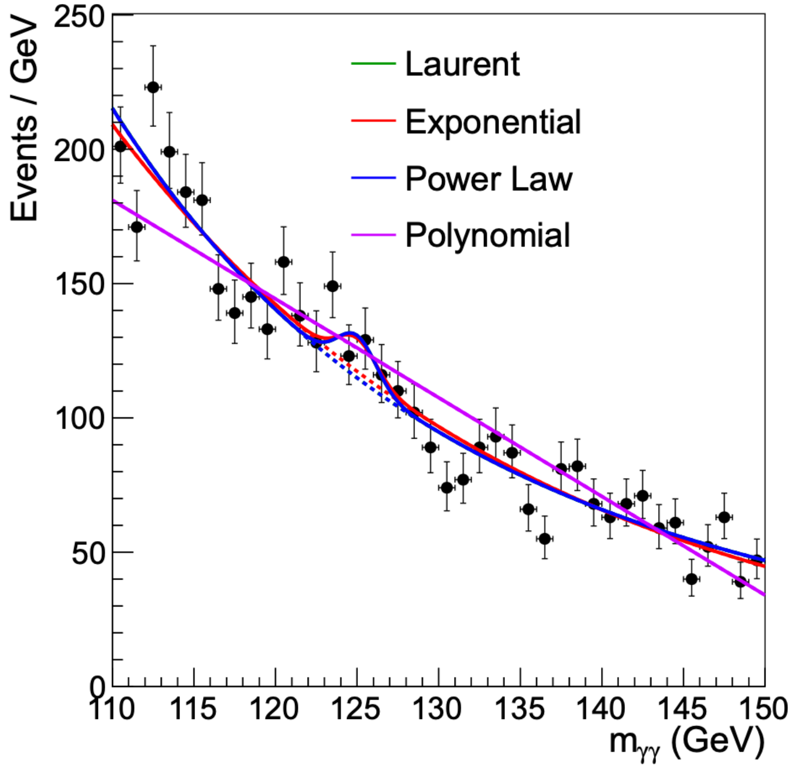
→ E.G tt+HF background in ttH(→bb) measurements includes uncertainties from comparing treatment of *b*-quark in parton density functions

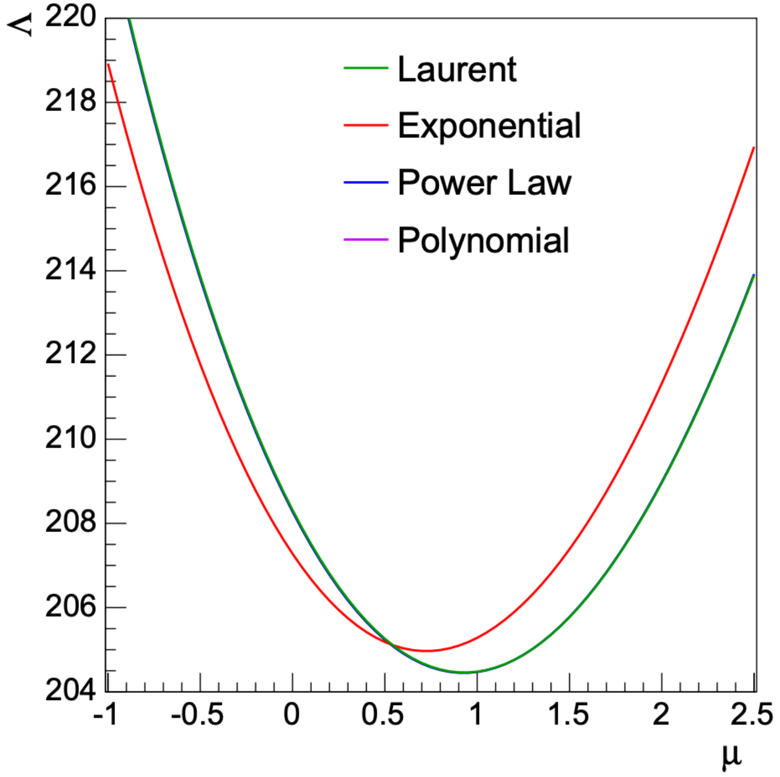
LHCHWG-2022-003



Background function choice

Ad-hoc parameterization of background function for peak-fitting is a source of uncertainty
→ Without clear motivation for one parameter or another, different choices result in different measured values of parameter of interest





Background function choice

Include potential bias from using wrong function as additional source of uncertainty in likelihood → parameterised as a Gaussian constrained nuisance parameter

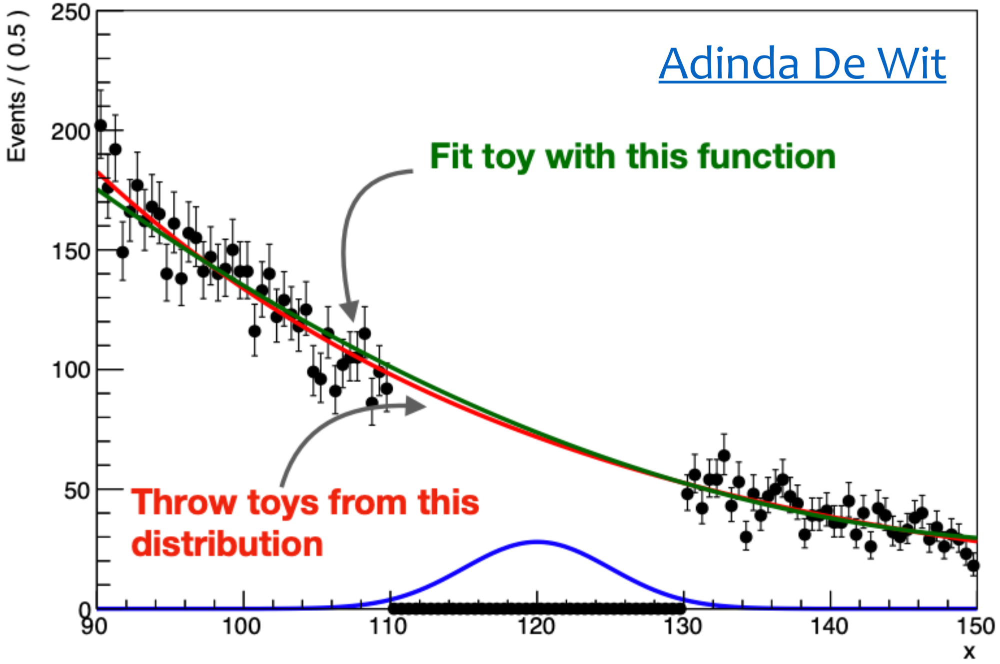
- Generate toys from the **red (”true”) function** with known **signal** strength
- Fit those toys with the **green background function** + measuring the amount of **signal**
- Bias in the amount of fitted signal induced by fitting with the green function instead of the red → distribution of fitted signal - injected signal
- Include as additional component when measuring signal

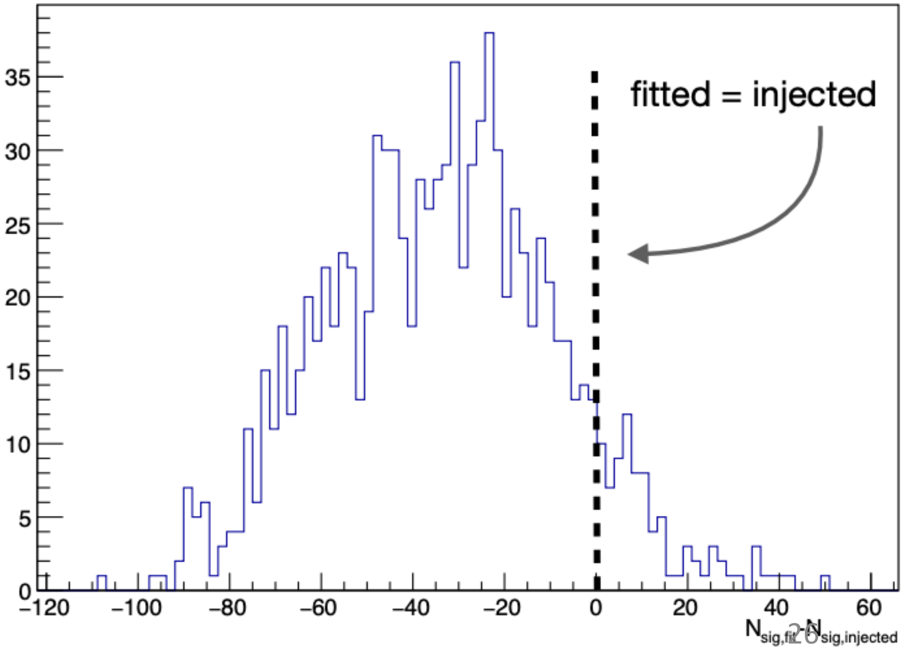
$$\mu \cdot s_i(\vec{\theta}) \rightarrow \mu \cdot s_i(\vec{\theta}) + \text{bias} \cdot \theta_{\text{bias}}$$

N_{events} bias

Unit gaussian

- Alternative approach (discrete profiling): [JINST (2015) 10 P04015]





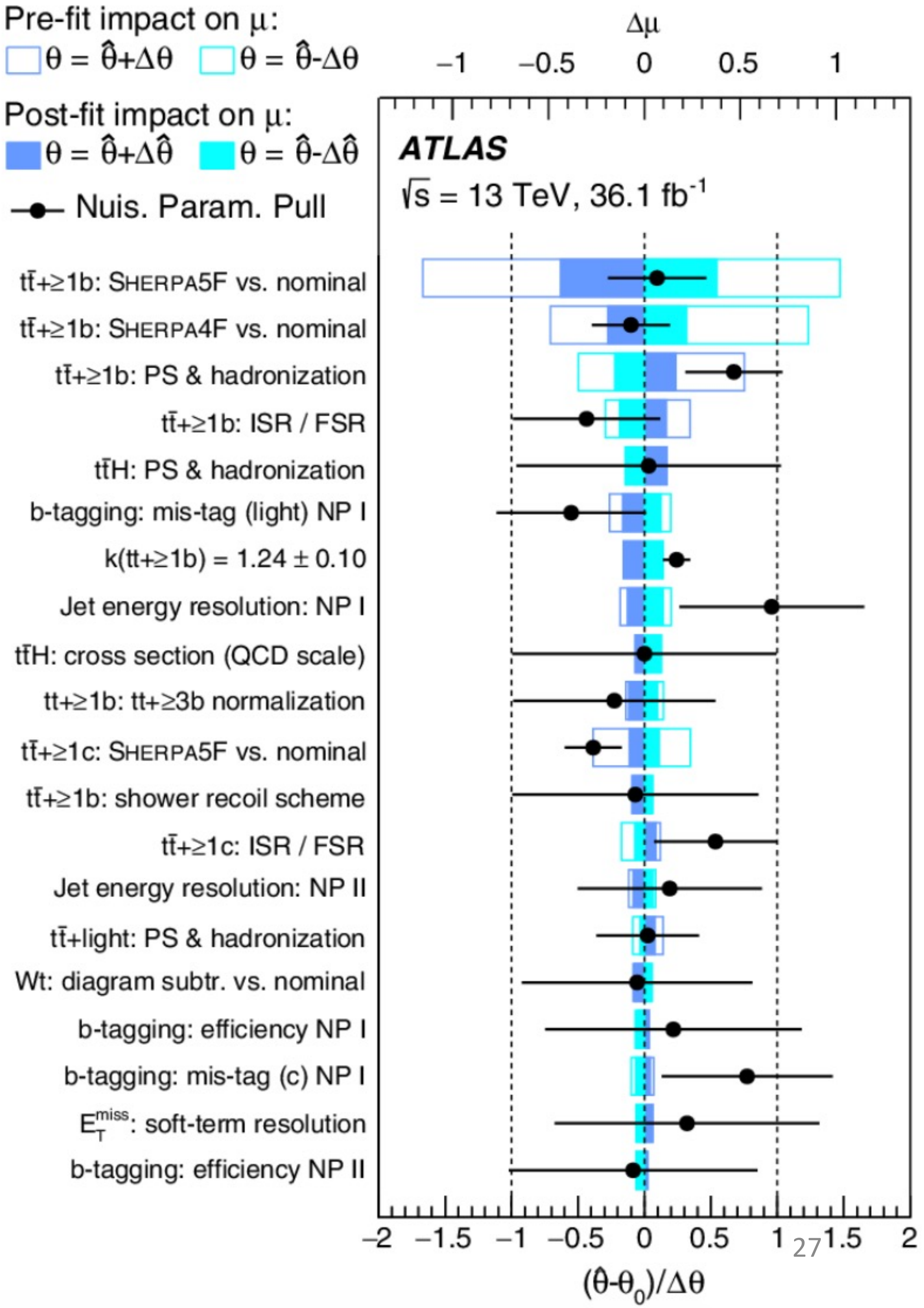
How do we know what matters?

Diagnostic tools help to quantify which sources of systematic uncertainty are most relevant in particular measurement

→ e.g measurement "**impact**" defined as

$$\Delta(\mu) = \hat{\mu}(\theta = \hat{\theta} + \Delta\theta) - \hat{\mu}(\theta = \hat{\theta})$$

→ Note similarity to OPAT → nuisance parameters can be varied one at a time to assess effect on measurement



How do we know what matters?

Diagnostic tools help to quantify which sources of systematic uncertainty are most relevant in particular measurement

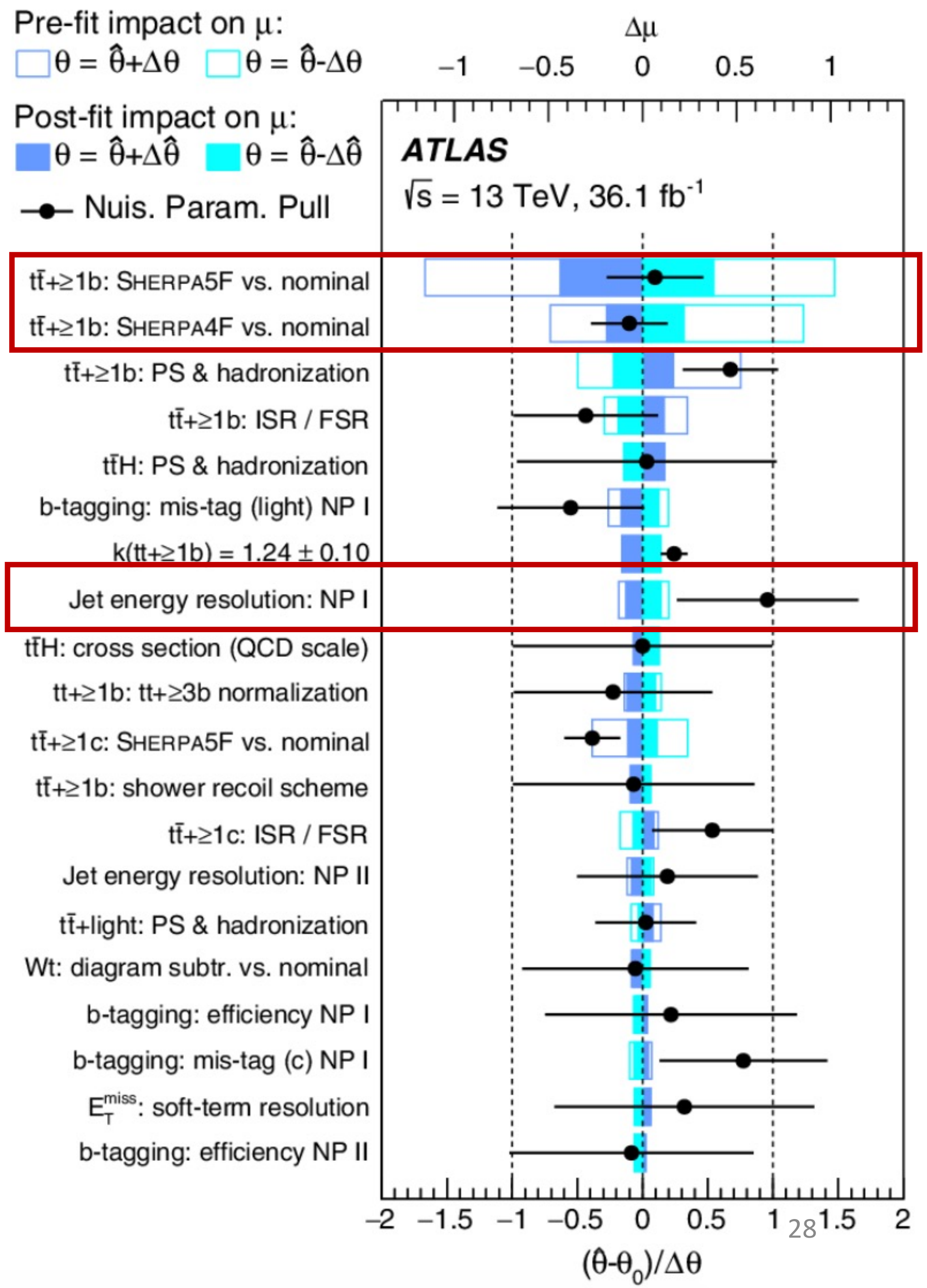
→ e.g measurement "**impact**" defined as

$$\Delta(\mu) = \hat{\mu}(\theta = \hat{\theta} + \Delta\theta) - \hat{\mu}(\theta = \hat{\theta})$$

→ Note similarity to OPAT → nuisance parameters can be varied one at a time to assess effect on measurement

Discussion points:

- *"Are we ok that our tt analysis fit updates our knowledge of **Jet Energy Resolution**?"*
- *"Is there a better way to include uncertainty due to model choice **than taking difference between (eg) simulations**?"*



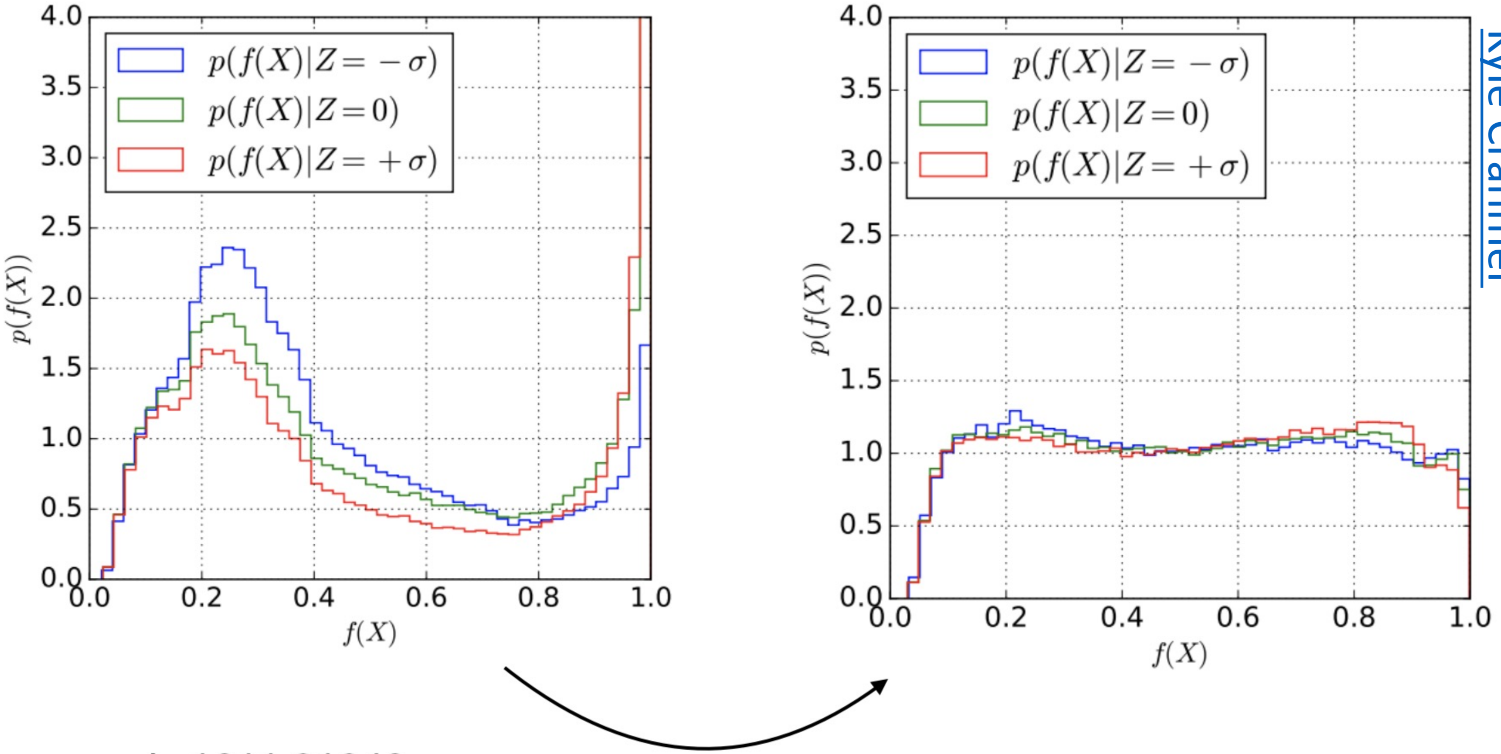
Can we avoid what doesn’t matter all together?

Machine learning methods to “learn” the variation of probability density wrt systematic variations

→ E.g Use adversarial methods to teach ML classifiers to reduce effect of systematic uncertainties on observable distribution!

→ See example : recent CMS LLP tagger using this approach ["A deep neural network to search for new long-lived particles decaying to jets"]()

→ Lots of overlap with discussions around LFI/SBI [1] for PP applications (see session this afternoon)



arxiv:1611.01046

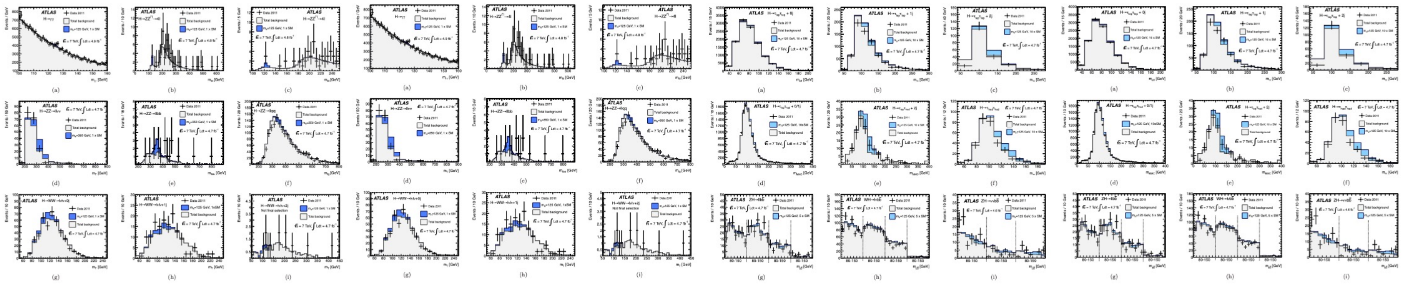
adversarial training

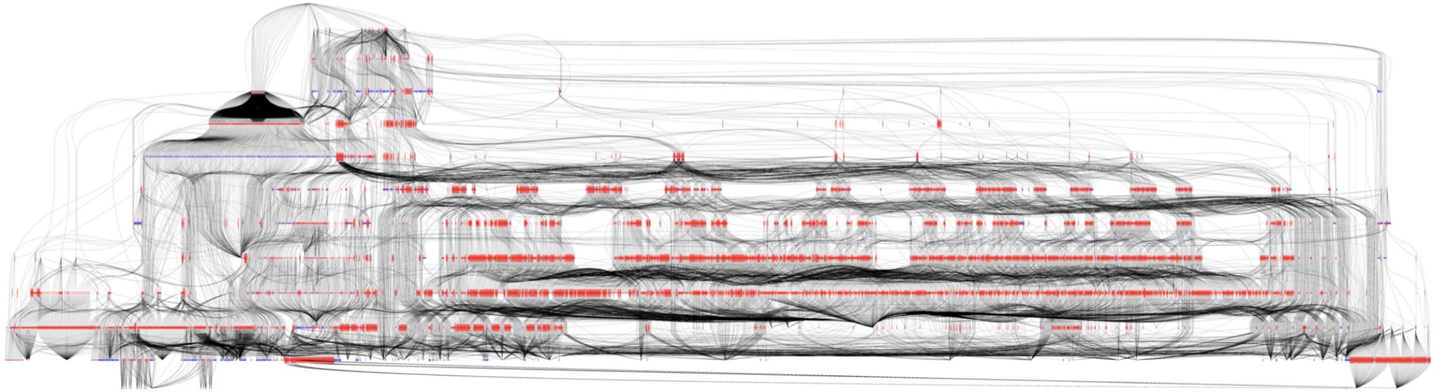
Kyle Cranmer

[1] See arXiv:1907.10621 for a good review

Communicating what we did

Particle physics experiments involve extremely complex likelihoods (many different regions in data and many hundreds of nuisance parameters)
→ Communicating this to statisticians & preservation of statistical model a challenge!

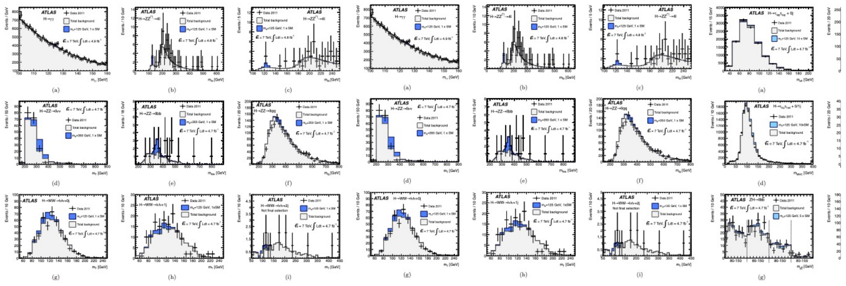


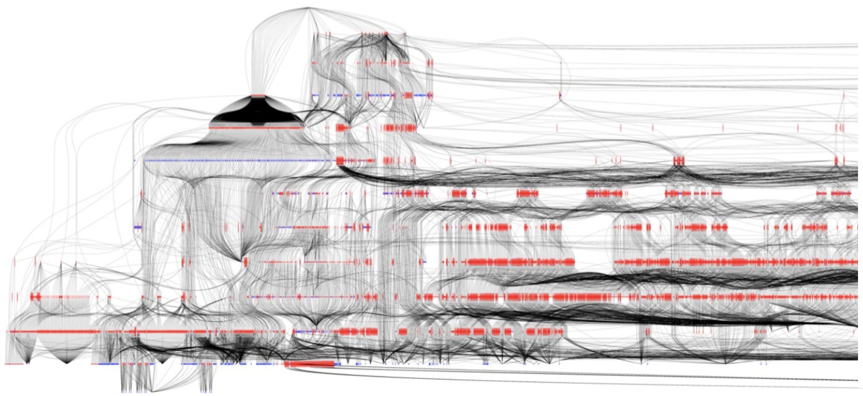


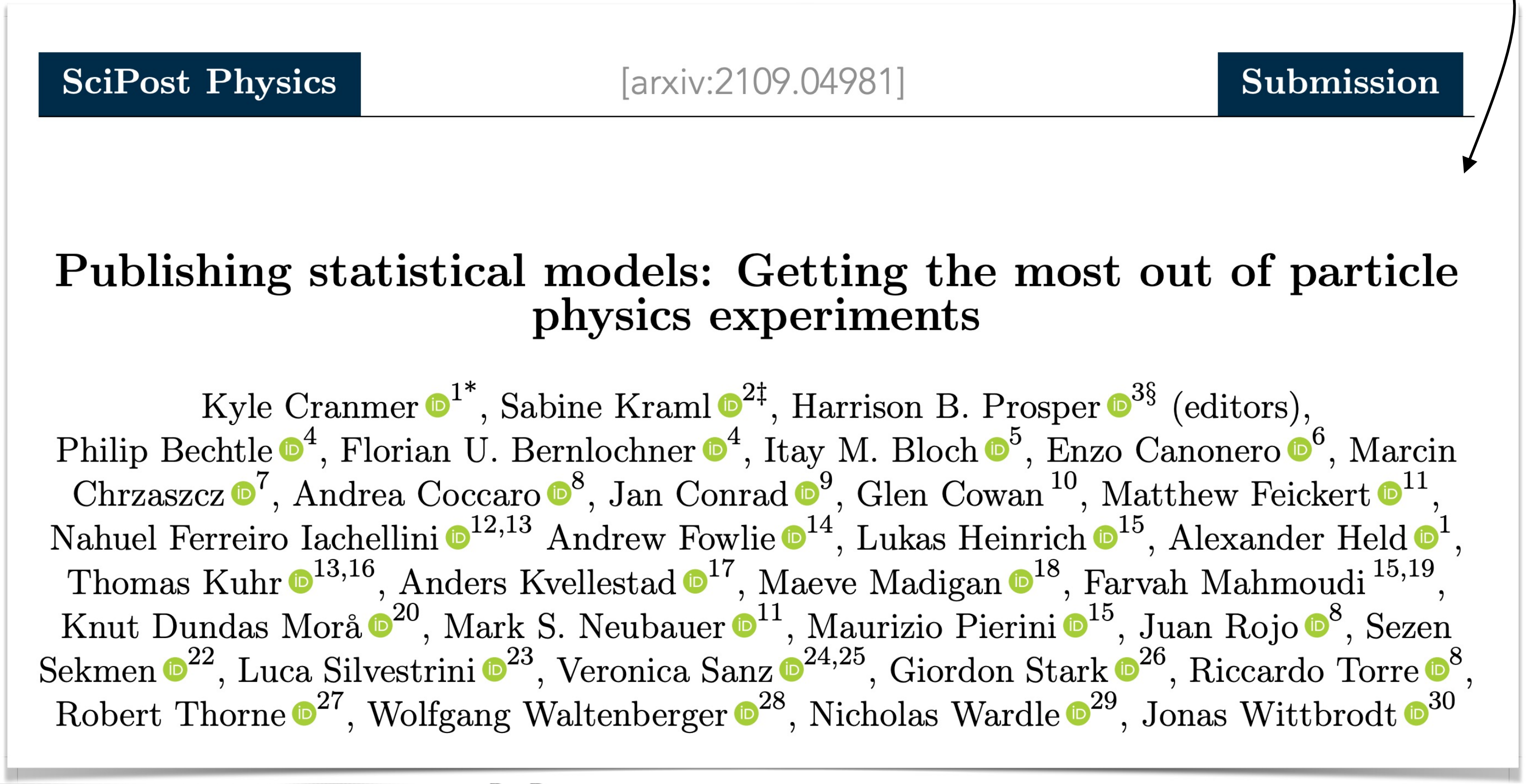
Communicating what we did

Particle physics experiments involve extremely complex likelihoods (many different regions in data and many hundreds of nuisance parameters)
→ Communicating this to statisticians & preservation of statistical model a challenge!

Recent paper to (finally) push for publication of these models!







SciPost Physics [arxiv:2109.04981] Submission

Publishing statistical models: Getting the most out of particle physics experiments

Kyle Cranmer[1*], Sabine Kraml[2‡], Harrison B. Prosper[3§] (editors), Philip Bechtle[4], Florian U. Bernlochner[4], Itay M. Bloch[5], Enzo Canonero[6], Marcin Chrzaszcz[7], Andrea Coccaro[8], Jan Conrad[9], Glen Cowan[10], Matthew Feickert[11], Nahuel Ferreiro Iachellini[12,13] Andrew Fowlie[14], Lukas Heinrich[15], Alexander Held[1], Thomas Kuhr[13,16], Anders Kvellestad[17], Maeve Madigan[18], Farvah Mahmoudi[15,19], Knut Dundas Morå[20], Mark S. Neubauer[11], Maurizio Pierini[15], Juan Rojo[8], Sezen Sekmen[22], Luca Silvestrini[23], Veronica Sanz[24,25], Giordon Stark[26], Riccardo Torre[8], Robert Thorne[27], Wolfgang Waltenberger[28], Nicholas Wardle[29], Jonas Wittbrodt[30]

SciPost Phys. 12, 037 (2022)

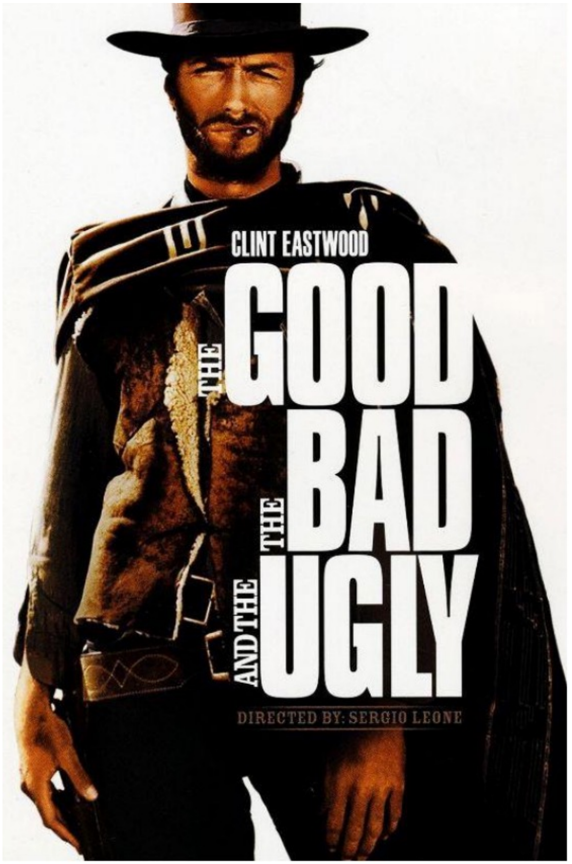
Summary of my Summary

Particle physics experiments deal with a **vast array of sources of systematic uncertainties**

→ A particle physicist might spend most of their time on dealing with these for a publication!

Discussion items (or Q's for a statistician):

- When reporting uncertainties in OPAT is providing covariance enough?
- When modelling systematic uncertainties using nuisance parameters, Should we sample many different parameter values to build suitable parameterisation and are there smarter ways (eg GPs/ML?) to automate this?
- Are we ok that our fit updates our knowledge of certainty nuisance parameters?
- Is there a better way to include uncertainties due to model choice than taking difference between (eg) simulations? Or approaches such as inflating uncertainty to cover potential bias / discrete profile method?



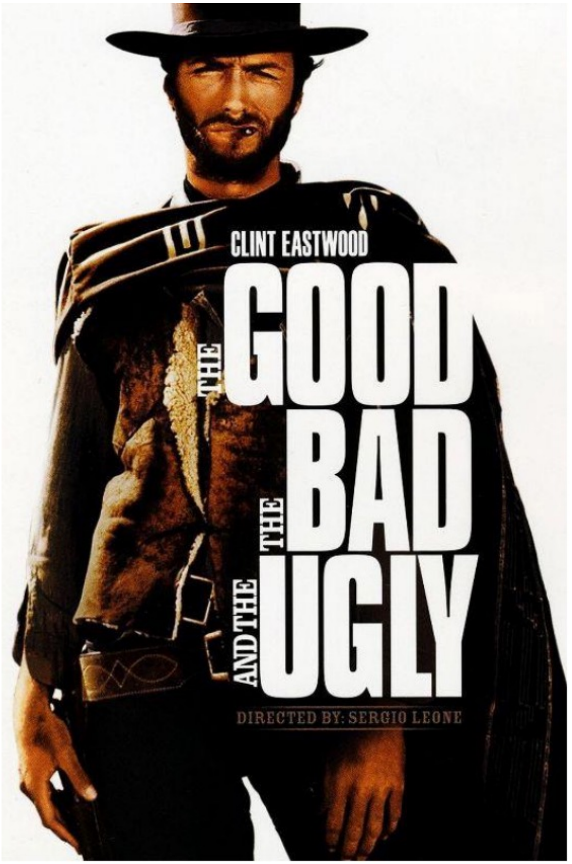
Summary of my Summary

Particle physics experiments deal with a **vast array of sources of systematic uncertainties**

→ A particle physicist might spend most of their time on dealing with these for a publication!

Discussion items (or Q's for a statistician):

- When reporting uncertainties in OPAT is providing covariance enough?
- When modelling systematic uncertainties using nuisance parameters, Should we sample many different parameter values to build suitable parameterisation and are there smarter ways (eg GPs/ML?) to automate this?
- Are we ok that our fit updates our knowledge of certainty nuisance parameters?
- Is there a better way to include uncertainties due to model choice than taking difference between (eg) simulations? Or approaches such as inflating uncertainty to cover potential bias / discrete profile method?



Thanks for your Attention!

Backup Slides

Errors on Errors?

[Eur. Phys. J. C (2019) 79:133]

Uncertainty on parameterization model due to limited MC samples could be addressed by including errors on the parameterization (errors on errors)

Statistical Models with Uncertain Error Parameters

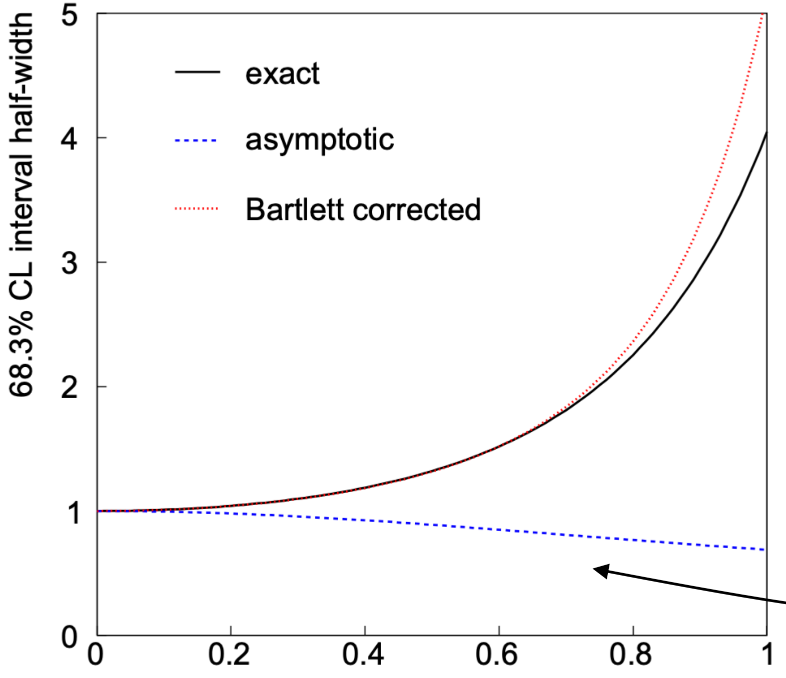
Glen Cowan

Physics Department, Royal Holloway, University of London, Egham, TW20 0EX, U.K.

Received: date / Revised version: date

Abstract. In a statistical analysis in Particle Physics, nuisance parameters can be introduced to take into account various types of systematic uncertainties. The best estimate of such a parameter is often modeled as a Gaussian distributed variable with a given standard deviation (the corresponding "systematic error"). Although the assigned systematic errors are usually treated as constants, in general they are themselves uncertain. A type of model is presented where the uncertainty in the assigned systematic errors is taken into account. Estimates of the systematic variances are modeled as gamma distributed random variables. The resulting confidence intervals show interesting and useful properties. For example, when averaging measurements to estimate their mean, the size of the confidence interval increases for decreasing goodness-of-fit, and averages have reduced sensitivity to outliers. The basic properties of the model are presented and several examples relevant for Particle Physics are explored.

PACS. 02.50.Tt Inference methods – 02.70.Rr General statistical methods



Typical test-statistics can be modified to maintain good asymptotic properties

[Alessandra Brazzale]

r = relative error on error

Theory uncertainties

Frank Tackmann

- The *theory uncertainty* is due to the fact that in many cases the formula itself is not fully exact (e.g. derived in some approximation)
  - It is *not* the inexact knowledge of parameters needed in the (otherwise exact) formula (like the length of the pendulum)

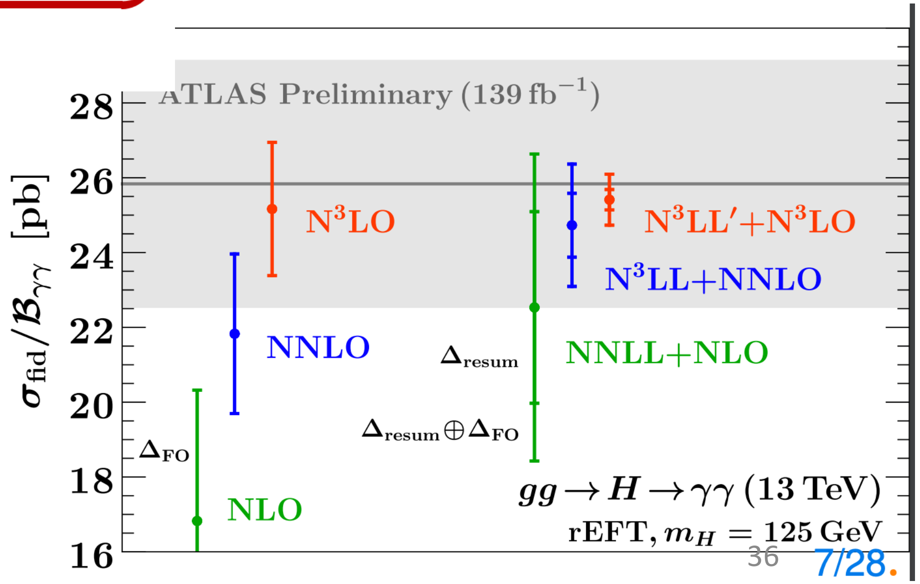
$$\sigma = c_0 + \alpha(\mu_0)\, c_1 + \alpha^2(\mu_0)\, c_2 + \cdots$$

$$= c_0 + \alpha(\mu)\, c_1 + \underbrace{\alpha^2(\mu)(c_1\, b_0 \ln \mu/\mu_0 + c_2) + \cdots}_{\text{neglected}}$$

$$= \sigma \pm \Delta\sigma$$

Aim to estimate uncertainty due to contribution of **neglected terms**

- Typical strategy of using "scale variations" not always guaranteed to cover most accurate calculation
- Not obvious what to assume for distribution – is Δσ the width of a Gaussian?
- How to correlate different sources (scale parameters are not real parameters of the model)



Theory uncertainties

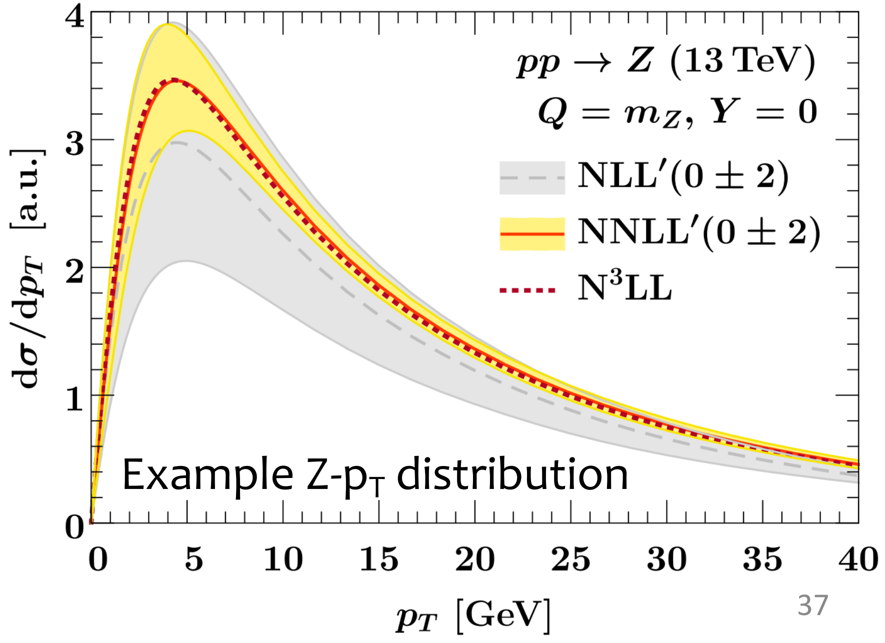
Frank Tackmann

- The *theory uncertainty* is due to the fact that in many cases the formula itself is not fully exact (e.g. derived in some approximation)
  - It is *not* the inexact knowledge of parameters needed in the (otherwise exact) formula (like the length of the pendulum)

$$\sigma = c_0 + \alpha(\mu_0)\, c_1 + \alpha^2(\mu_0)\, c_2 + \cdots$$

Instead identify the actual source of uncertainty and parameterize our knowledge of the structure of missing terms

- Introduce “Theory nuisance parameters” → genuine parameterization of missing terms
- Appeal to CLT for total theory uncertainty → Gaussian distribution well motivated



Example Z-p_T distribution

LIKELIHOODS

SEARCH DATA CALIBRATION OTHER MEASUREMENTS /CONSTRAINTS

$$\mathcal{L}(s, \vec{\theta}_s, \vec{\theta}_b) = \mathcal{L}_{\text{sci}}(s, \vec{\theta}_s, \vec{\theta}_b) \times \mathcal{L}_{\text{cal}}(\vec{\theta}_b) \times \mathcal{L}_{\text{anc}}(\vec{\theta}_b)$$

COUNTING

$$\mathcal{L}_{\text{sci}}(s, \vec{\theta}_s, \vec{\theta}_b) = \text{Poisson}(N_{\text{sci}} \,|\, \mu_b(\vec{\theta}_b) + \mu_s(s, \vec{\theta}_s, \vec{\theta}_b))$$

ON-OFF LIKELIHOODS

$$\mathcal{L}_{\text{sci}}(s, \vec{\theta}_s, \vec{\theta}_b) = \text{Poisson}(N_{\text{sci}} \,|\, \mu_b(\vec{\theta}_b) + \mu_s(s, \vec{\theta}_s, \vec{\theta}_b)) \times \text{Poisson}(N_{\text{cal}} \,|\, \alpha \times \mu_b(\vec{\theta}_b))$$

BINNED LIKELIHOODS

$$\mathcal{L}_{\text{sci}}(s, \vec{\theta}_s, \vec{\theta}_b) = \prod_{i=1}^{N_s} \left[\text{Poisson}(N_i \,|\, \mu_{b,i}(\vec{\theta}_b) + \mu_{s,i}(s, \vec{\theta}_s, \vec{\theta}_b)) \right]$$

Knut Moraa

UNBINNED LIKELIHOODS

$$\mathcal{L}_{\text{sci}}(s, \vec{\theta}_s, \vec{\theta}_b) = \text{Poisson}(N_{\text{sci}} \,|\, \mu_b(\vec{\theta}_b) + \mu_s(s, \vec{\theta}_s, \vec{\theta}_b)) \times \prod_{i=1}^{N_s} \left[\frac{\mu_s}{\mu_s + \mu_b} f_s(\vec{x}_i \,|\, s, \vec{\theta}_s, \vec{\theta}_b) + \frac{\mu_b}{\mu_s + \mu_b} f_b(\vec{x}_i \,|\, \vec{\theta}_b) \right]$$

$\mathcal{L}_{\text{cal}}(\vec{\theta}_b)$ typically on the same form, while $\mathcal{L}_{\text{anc}}(\vec{\theta}_b)$ contains ancillary measurements— often Gaussian terms like $\text{Gaussian}(\hat{\theta}_i \,|\, \theta_i, \sigma_{\theta_i})$ but sometimes more complex functions, e.g. with correlations or with a different likelihood shape

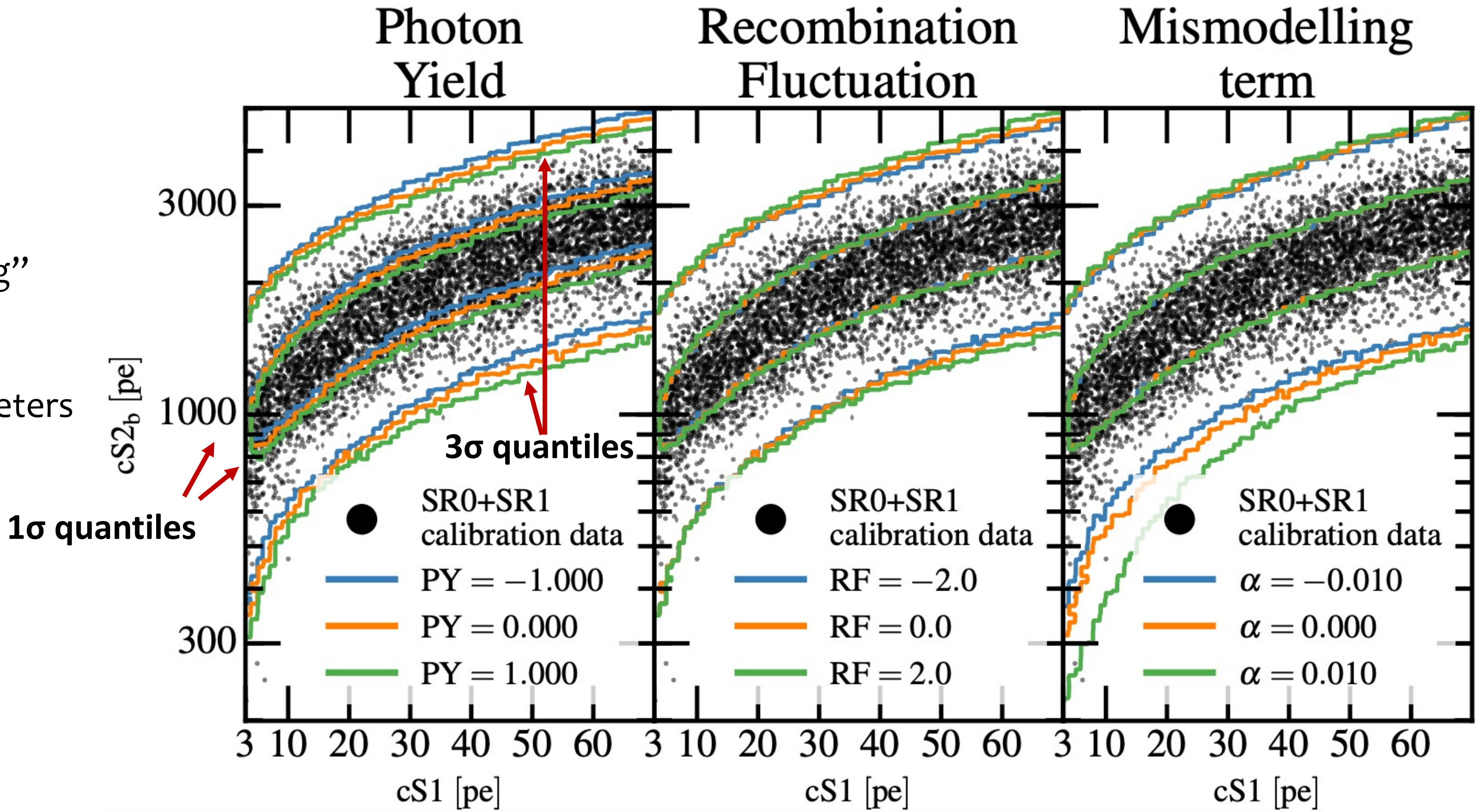
Template morphing

Knut Moraa

Multiple bins in one or more observables (shape) analysis requires “template morphing”

E.G Rn220 calibration data: electron-recoil model parameters varied

→ How to map systematic variations in predicted distribution onto a nuisance parameter?

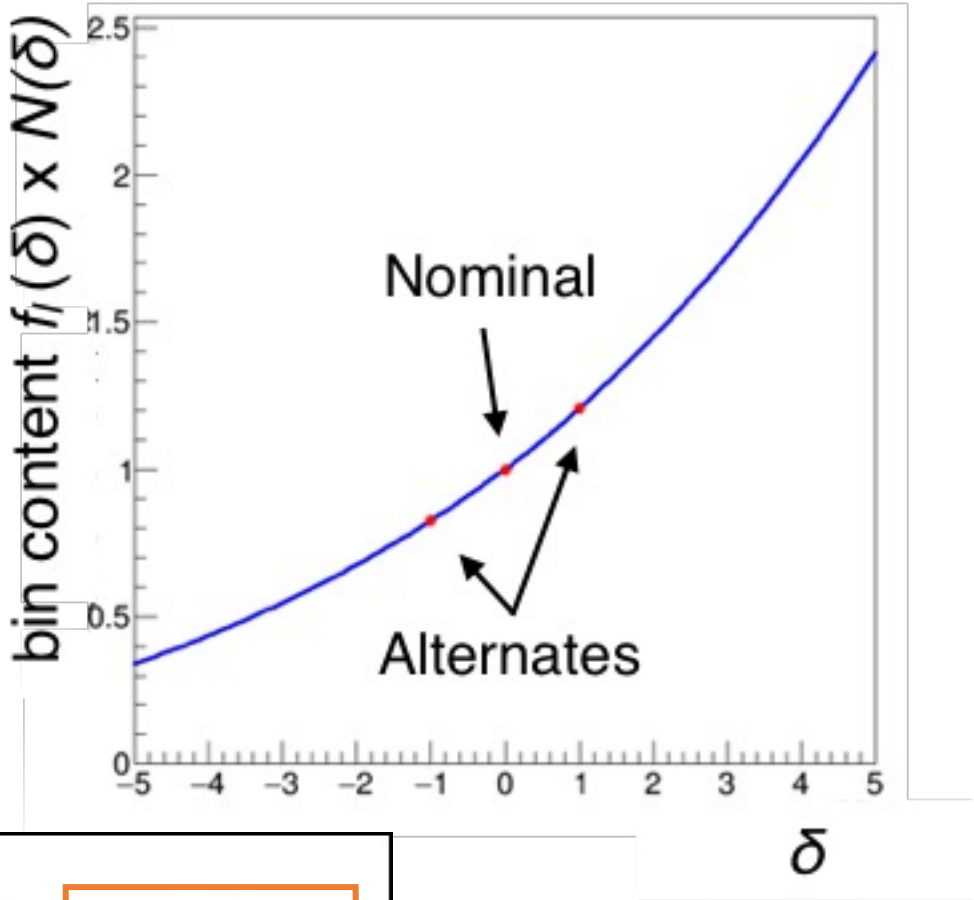


Shape interpolations

The effects of correlated systematic uncertainties on n_I are modelled using quadratic(linear) interpo(extrapo)lation function

$$\boxed{f_I(\boldsymbol{\delta}) = f_I^0 \cdot \frac{1}{F(\boldsymbol{\delta})} \prod_j p_{Ij}(\delta_j)}$$

$$F(\boldsymbol{\delta}) = \sum_I f_I(\boldsymbol{\delta})$$

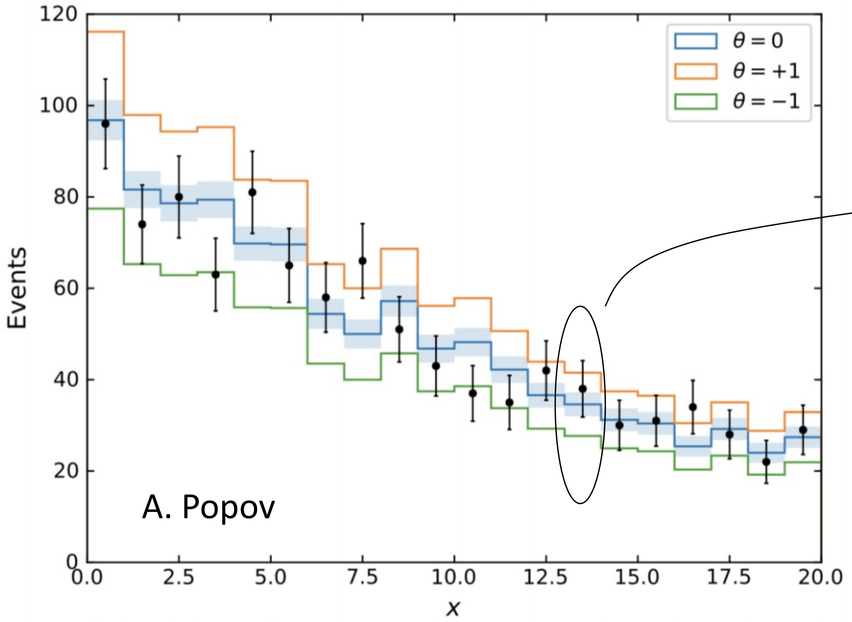


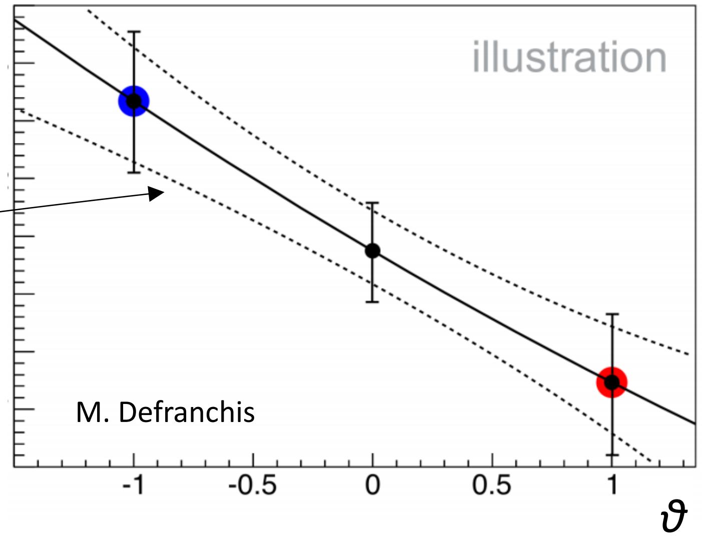
$$p_{Ij}(\delta_j) = \begin{cases} \frac{1}{2}\delta_j(\delta_j - 1)\kappa_{Ij}^- - (\delta_j - 1)(\delta_j + 1) + \frac{1}{2}\delta_j(\delta_j + 1)\kappa_{Ij}^+ & \text{for } |\delta_j| < 1 \\ \left[\frac{1}{2}(3\kappa_{Ij}^+ + \kappa_{Ij}^-) - 2\right]\delta_j - \frac{1}{2}(\kappa_{Ij}^+ + \kappa_{Ij}^-) + 2 & \text{for } \delta_j > 1 \\ \left[2 - \frac{1}{2}(3\kappa_{Ij}^- + \kappa_{Ij}^+)\right]\delta_j - \frac{1}{2}(\kappa_{Ij}^+ + \kappa_{Ij}^-) + 2 & \text{for } \delta_j < -1 \end{cases}$$

Caveats for shape interpolations!

Procedure relies on ***smooth*** templates to extract polynomial coefficients

→ Limited sample sizes in Monte Carlo *or* data used to determine alternative templates can lead to unphysical shapes !





How to account for uncertainties in alternative templates?

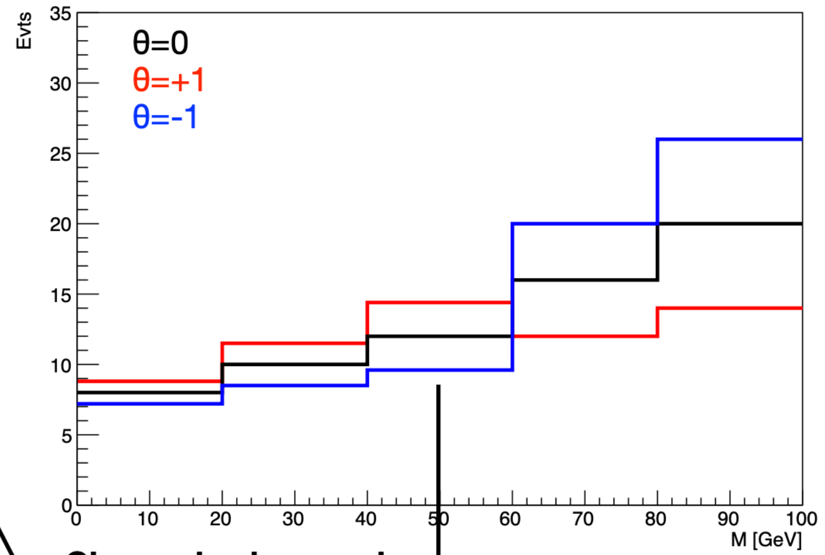
→ See G. Cowan [Eur. Phys. J. C (2019) 79:133](#) for more "uncertainties on uncertainties"

$$\boxed{\frac{1}{F(\boldsymbol{\delta})}\prod_j p_{Ij}(\delta_j)}$$

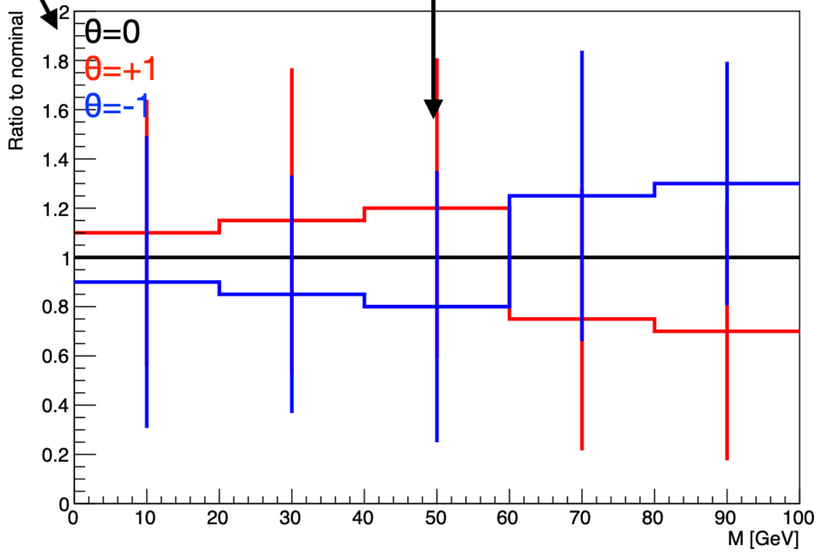
2. This approach *assumes* we can *factorize* effects on the shape from different parameters

On shape-changing systematics

- Are shape-changing effects **genuine**?
  - Don't want to model "noise"
- Consider the example given earlier
  - Looks like the changes are large!
- Now let's also look at the uncertainties



Change in shape only



Background function choice

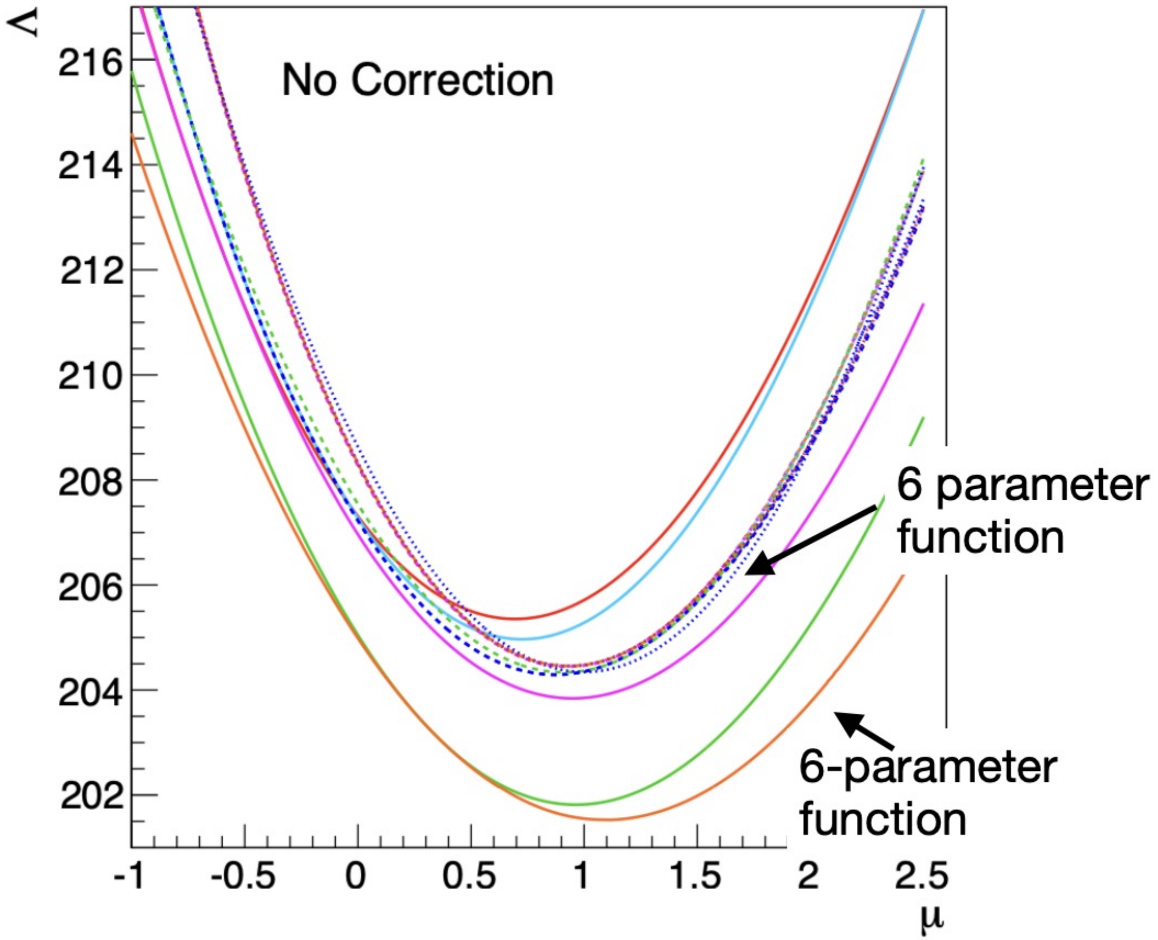
Adinda De Wit

Alternative - treat choice of functional form as a nuisance parameter in the fit

- Label each function with a discrete index and treat that index as a nuisance parameter
- Minimise likelihood across all choices of index to obtain profiled likelihood
- Correct for # parameters -> $-2\Delta\log L \rightarrow -2\Delta\log L + cN_{pars}$

Requires detailed studies of coverage and tune of "c" parameter

https://arxiv.org/abs/1408.6865 (JINST)



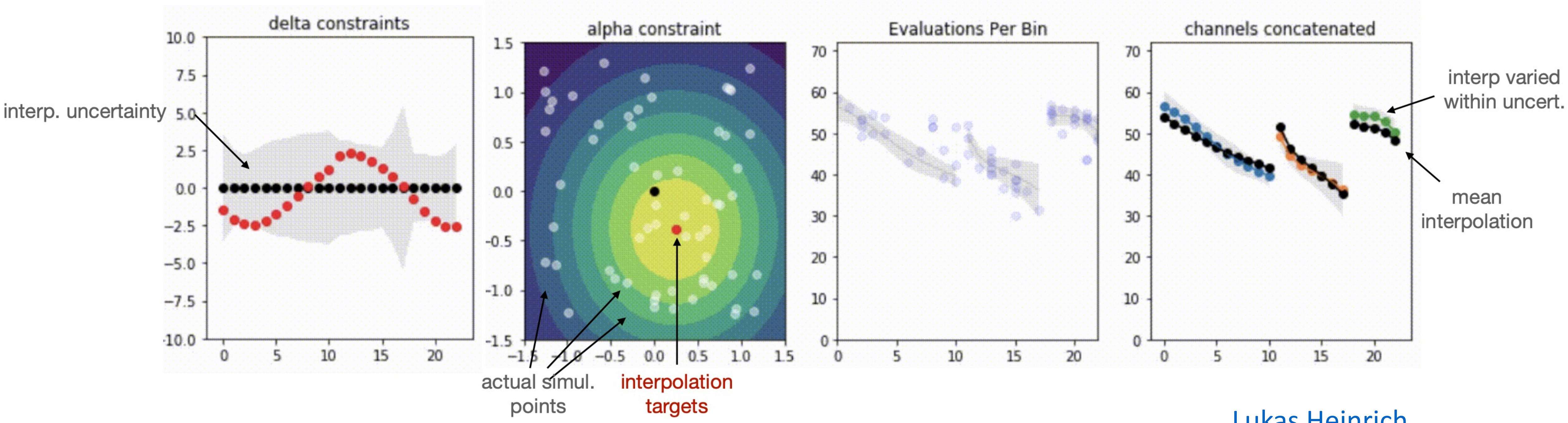
Smarter interpolation?

Gaussian processes offer an appealing solution to this interpolation problem

→ Particularly nice that Bayesian picture offers route to estimate uncertainty on interpolation itself (see later errors on errors!)

→ Other interpolation thoughts in Tudor Manole's talk

Example: Template Interpolation base on variable set of input templates



Lukas Heinrich

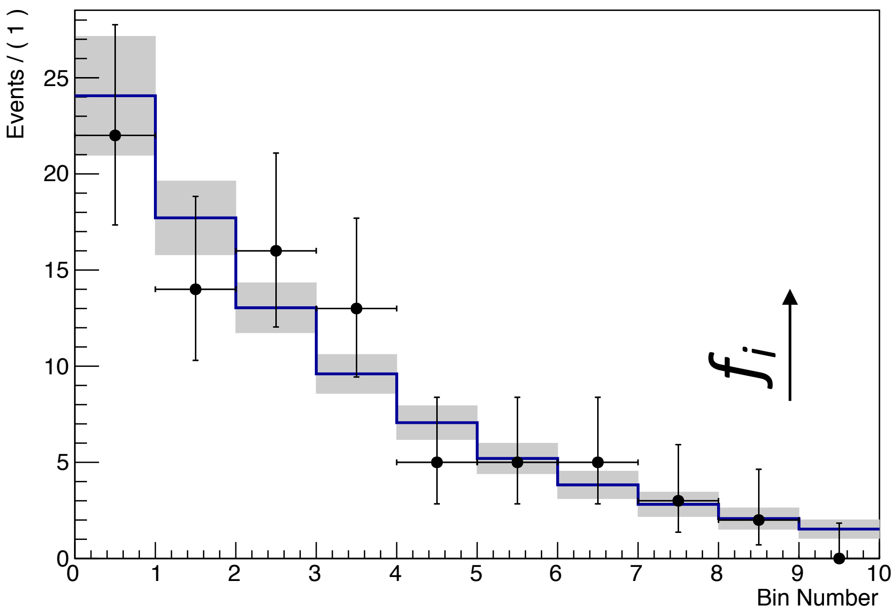
Goodness of Fit (toy example)

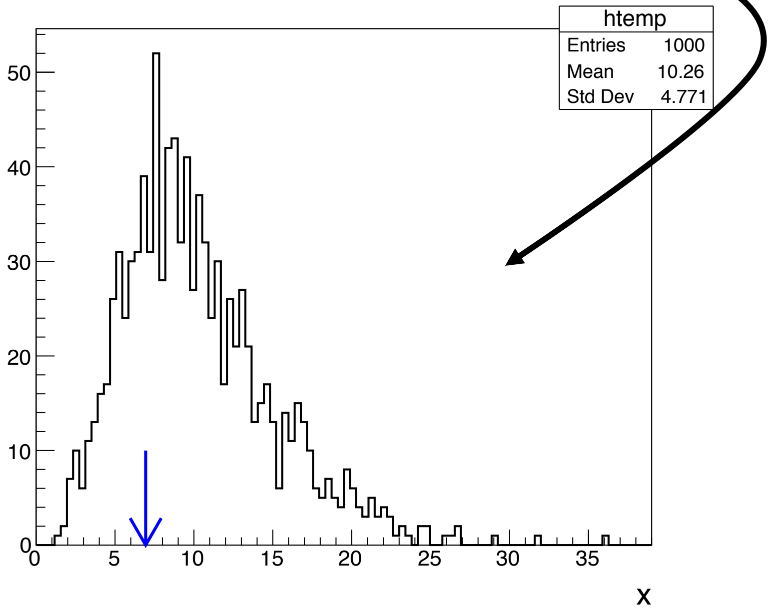
Compatibility of data and prediction in distributions used for combined measurements –
No alternate hypothesis → how to make a likelihood test?

Make use of the **saturated likelihood** : compare likelihood at “best-fit” under model with that of best possible fit to data …

$$x = 2\sum_{i} f_i(\hat{\boldsymbol{\theta}}) - d_i + d_i \ln \frac{d_i}{f_i(\hat{\boldsymbol{\theta}})}$$

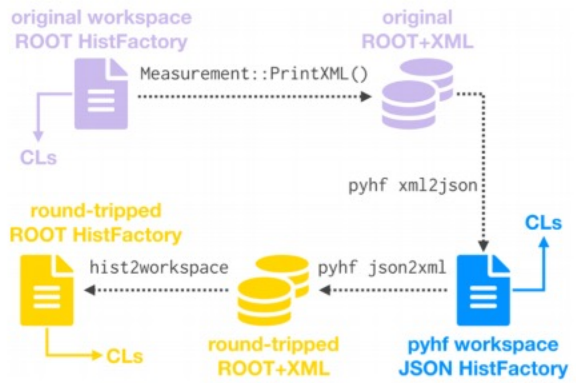
Generate toys under f_i to calculate p-value





Published likelihoods

ATLAS-PHYS-PUB-2019-029



Can publish full likelihoods – eg.using pyHF

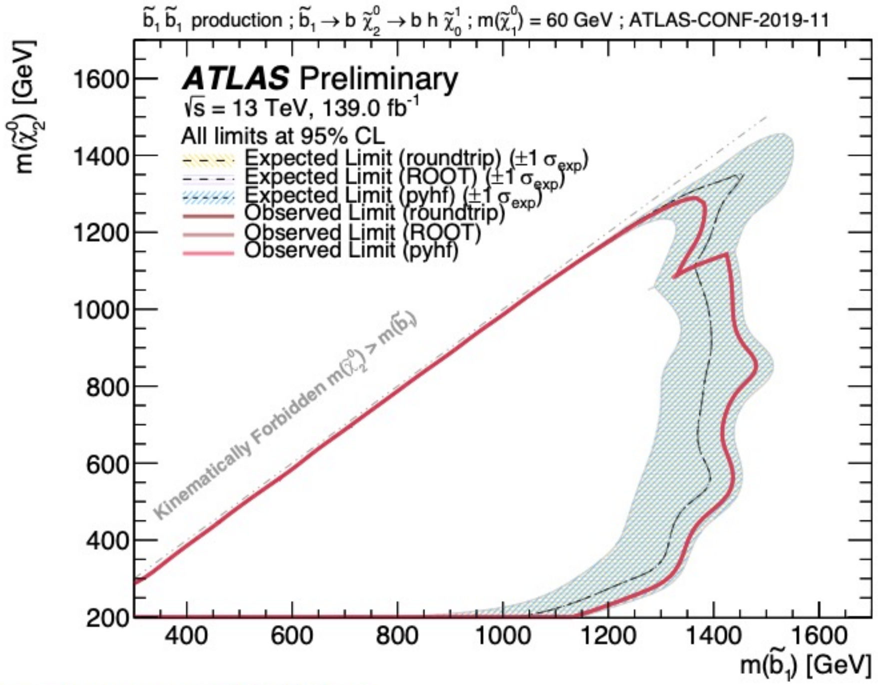
JSON based (ROOT/XML free) encoded workspaces containing ***full*** likelihood model

Example: ATLAS search for sbottom production

Likelihood validated against ROOT based (experimental) results by comparing exclusion contours

Can swap out components of likelihood directly inside workspaces!

→ With care, one can build joint (**combined**) likelihoods from these inputs



Similar tool HepLike : https://github.com/mchrzasz/HEPLike