Robust estimates of theoretical uncertainties at fixed-order in perturbation theory

Rene Poncelet

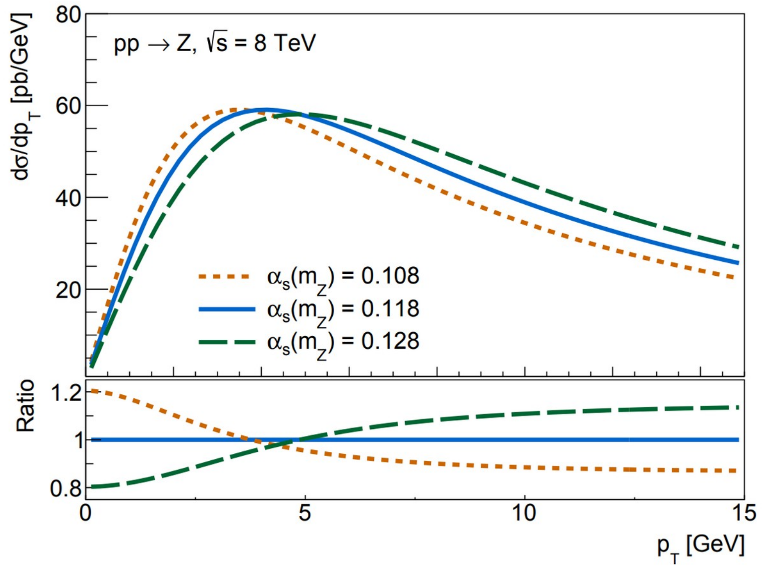
based on [Lim, Poncelet, 2412.14910]

THE HENRYK NIEWODNICZAŃSKI
INSTITUTE OF NUCLEAR PHYSICS
POLISH ACADEMY OF SCIENCES

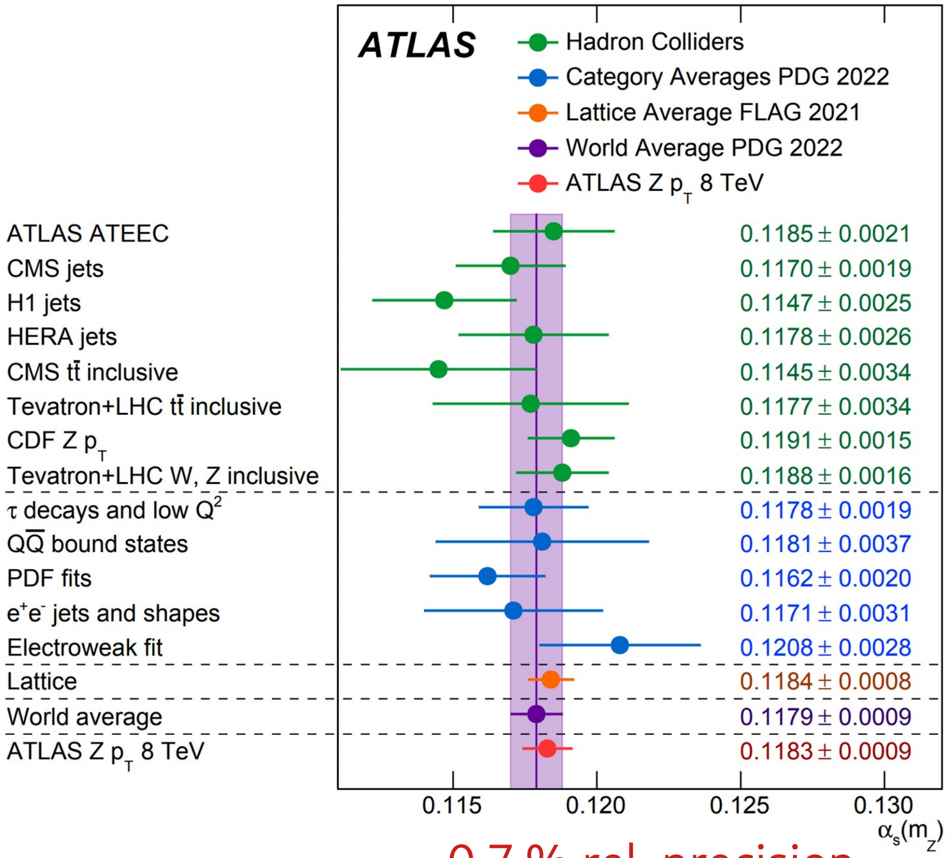
Precision example: strong coupling from pT(Z)

[ATLAS 2309.12986]

Sensitivity of Z-boson's recoil to the strong coupling constant:



→ at low pT resummation regime!
→ theory uncertainty?



~0.7 % rel. precision

Uncertainties in precision phenomenology

Experiments are getting more precise → theory uncertainties matter!

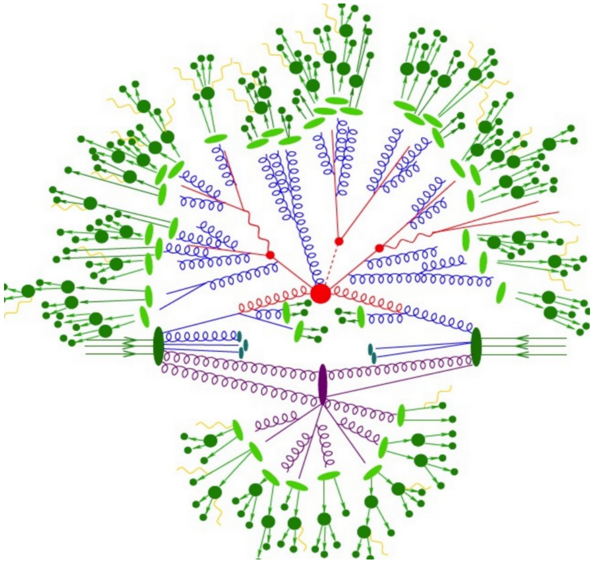
Sources of theory uncertainties:

- parametric (values of coupling parameters etc.)
→ variation of parameters within their uncertainties

- parton distribution functions (PDFs)
→ different error propagation methods (fit parameter, replicas,...)

- non-perturbative parameters in Monte Carlo simulations.
→ needs data constraints by definition. Problematic if dominant effect...

- missing higher orders in fixed-order and resummed predictions (MHOU)
→ tricky because we are trying to estimate the unknown….



[Credit: SHERPA]

Missing higher orders

Notation from: [Tackmann 2411.18606]
Beyond Scale Variations: Perturbative Theory Uncertainties from Nuisance Parameters

Generic perturbative expansion:

$$f(\alpha) = f_0 + f_1\alpha + f_2\alpha^2 + f_3\alpha^3 + \dots$$

f_i : the coefficient of the series, potentially unknown

We can compute the truncated series: $\hat{f}_i$: the true value, i.e. a value we actually computed

$$f^{\mathrm{LO}}(\alpha) = \hat{f}_0 \qquad f^{\mathrm{NLO}}(\alpha) = \hat{f}_0 + \hat{f}_1\alpha \qquad f^{\mathrm{NNLO}}(\alpha) = \hat{f}_0 + \hat{f}_1\alpha + \hat{f}_2\alpha^2$$

The missing terms are the source of uncertainty.
(assume convergence → the first missing is the dominant one)

$$f^{\mathrm{LO+1}}(\alpha) = \hat{f}_0 + f_1\alpha \quad f^{\mathrm{NLO+1}}(\alpha) = \hat{f}_0 + \hat{f}_1\alpha + f_2\alpha^2 \quad f^{\mathrm{NNLO+1}}(\alpha) = \hat{f}_0 + \hat{f}_1\alpha + \hat{f}_2\alpha^2 + f_3\alpha^3$$

Challenge: how to estimate f_1, f_2, f_3, … without computing them?

Theory uncertainties from scale variations

Lets focus on QCD as an example: $\alpha = \alpha_s(\mu_0)$

$$\mathrm{d}\sigma = \mathrm{d}\sigma^{(0)} + \alpha_s \mathrm{d}\sigma^{(1)} + \alpha_s^2 \mathrm{d}\sigma^{(2)} + \dots \xrightarrow{\text{RGE}} \mu \frac{\mathrm{d}\sigma^{(n)}}{\mathrm{d}\mu} = \mathcal{O}\left(\alpha_s^{(n_0+n+1)}\right)$$

Of same order as the next dominant term → exploiting this to estimate size of $\mathrm{d}\sigma^{(n+1)}$

Scale variation prescription (ad-hoc and heuristic choice)

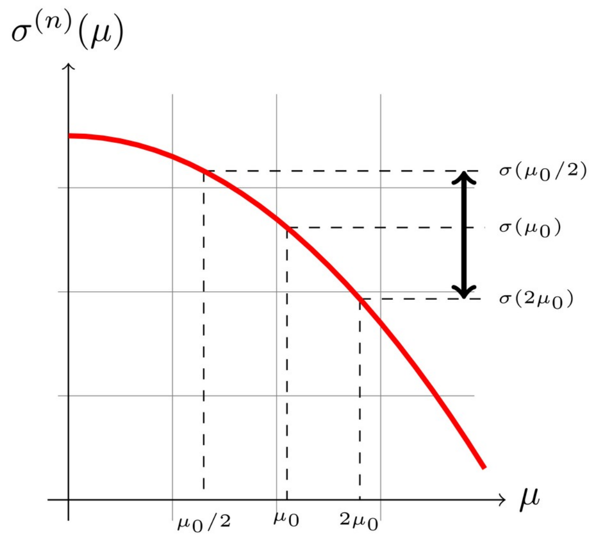
- choose 'sensible' μ_0

→ principle of fasted apparent convergence: $\sigma^{(n)}(\mu_{\mathrm{FAC}}) = 0$

→ principle of minimal sensitivity $\mu \frac{\mathrm{d}\sigma(\mu)}{\mathrm{d}\mu}\Big|_{\mu=\mu_{\mathrm{PMC}}} = 0$

→ ...

- vary with a factor (typically 2)
- take envelope as uncertainty



Scale variation approach

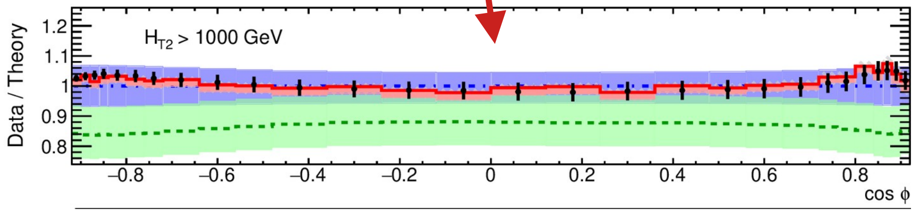
Estimates from scale variations for the unknown f_i:

$$f_2 \equiv \Delta f^{\mathrm{NLO}} = f^{\mathrm{NLO}} - \tilde{f}^{\mathrm{NLO}}(\tilde{\alpha}) = -\alpha^2 b_0 \hat{f}_1 + \mathcal{O}(\alpha^3)$$

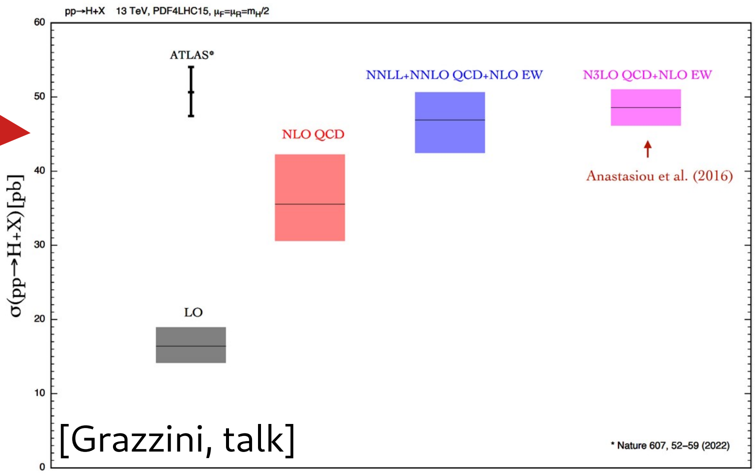
$$f_3 \equiv \Delta f^{\mathrm{NNLO}} = f^{\mathrm{NNLO}} - \tilde{f}^{\mathrm{NNLO}}(\tilde{\alpha}) = \alpha^3 (2b_0(\hat{f}_2 - b_0 \hat{f}_1) + b_1 \hat{f}_1) + \mathcal{O}(\alpha^4)$$

$$b_0 = \frac{\beta_0}{2\pi} L \qquad b_1 = \frac{\beta_0^2}{4\pi^2} L^2 + \frac{\beta_1}{8\pi^2} L \qquad b_2 = \frac{\beta_0^3}{8\pi^3} L^3 + \frac{5\beta_0\beta_1}{32\pi^2} L^2 + \frac{\beta_2}{32\pi^3} L \qquad L = \ln\frac{\mu_0}{\mu}$$

Sometimes it does works … sometimes not



[ATLAS 2301.09351]



[Grazzini, talk]

Short comings of scale variations

- not always reliable … however in most cases issues are understood/expected:
  new channels, phase space constraints, etc. → often we can design workarounds
- however, some issues are more fundamental:
  - → how to choose the central scale? → not a physical parameter, no 'true' value
    (Principle of fasted apparent convergence, principle of minimal sensitivity,…)
  - → how to propagate the estimated uncertainty, no statistical interpretation!
  - → what about correlations? Based on 'fixed form' of the lower orders and RGE.

(At the moment) two alternative approaches under investigation:

"Bayesian"

[Cacciari,Houdeau 1105.5152]
[Bonvini 2006.16293]
[Duhr, Huss, Mazeliauskas, Szafron 2106.04585]

"Theory Nuisance Parameter"

[Tackmann 2411.18606]
→ W mass extraction: [CMS 2412.13872]
[Cal,Lim, Scott,Tackmann Waalewijn 2408.13301]
[Lim, Poncelet, 2412.14910]

Introducing theory nuisance parameters (TNPs)

[Tackman 2411.18606]

Generic perturbative expansion:

$$f(\alpha) = f_0 + f_1\alpha + f_2\alpha^2 + f_3\alpha^3 + \ldots$$

$$f^{\mathrm{LO+1}}(\alpha) = \hat{f}_0 + f_1(\theta)\alpha$$

$$f^{\mathrm{NLO+1}}(\alpha) = \hat{f}_0 + \hat{f}_1\alpha + f_2(\theta)\alpha^2$$

$$f^{\mathrm{NNLO+1}}(\alpha) = \hat{f}_0 + \hat{f}_1\alpha + \hat{f}_2\alpha^2 + f_3(\theta)\alpha^3$$

Introduce a parametrisation of unknown coefficients in terms of

”Theory nuisance parameters” θ

Key features

- The parametrization such that there is a true value: $f_i(\hat{\theta}) = \hat{f}_i$

- Distributions of θ “known” (for example from already existing computations)

- “Expert knowledge” to construct such a parametrisation

Example: TNPs in resummed cross sections

[Tackman 2411.18606]

Transverse momentum resummation:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}p_T} = [\underbrace{H \times B_a \times B_b \times S}](\alpha_s, \ln\left(\frac{p_T}{Q}\right)) + \mathcal{O}\left(\frac{p_T}{Q}\right)$$

$$X(\alpha_s, L) = X(\alpha_s) \exp \int_0^L \mathrm{d}L' \{\Gamma(\alpha_s(L'))L' + \gamma_X(\alpha_s(L'))\}$$

Perturbative expansion of resummation ingredients:

$$X(\alpha_s) = X_0 + \alpha_s X_1 + \alpha_s^2 X_2 + \cdots \boxed{+ \alpha_s^n X_n(\theta)}$$

$$\Gamma(\alpha_s) = \alpha_s[\Gamma_0 + \alpha_s \Gamma_1 + \alpha_s^2 \Gamma_2 + \cdots \boxed{+ \alpha_s^n \Gamma_n(\theta)]}$$

$$\gamma(\alpha_s) = \alpha_s[\gamma_0 + \alpha_s \gamma_1 + \alpha_s^2 \gamma_2 + \cdots \boxed{+ \alpha_s^n \gamma_n(\theta)]}$$

Task: find suitable parametrization and variation range

These are numbers for simple processes → only need normalisation

TNP parametrisations for resummation

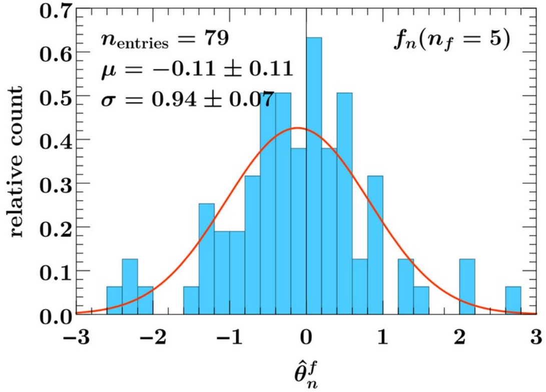
[Tackman 2411.18606]

$\gamma(\alpha_s)$	N_n	$\hat\gamma_0/N_0$	$\hat\gamma_1/N_1$	$\hat\gamma_2/N_2$	$\hat\gamma_3/N_3$	$\hat\gamma_4/N_4$
β	1	-15.3	-77.3	-362	-9652	-30941
	4^{n+1}	-3.83	-4.83	-5.65	-37.7	-30.2
	$4^{n+1}C_F C_A^n$	$\mathbf{-1.28}$	$\mathbf{-0.54}$	$\mathbf{-0.21}$	$\mathbf{-0.47}$	$\mathbf{-0.12}$
γ_m	1	-8.00	-112	-950	-5650	-85648
	4^{n+1}	-2.00	-7.028	-14.8	-22.1	-83.6
	$4^{n+1}C_F C_A^n$	$\mathbf{-1.50}$	$\mathbf{-1.76}$	$\mathbf{-1.24}$	$\mathbf{-0.61}$	$\mathbf{-0.77}$
$2\Gamma^q_{\rm cusp}$	1	$+10.7$	$+73.7$	$+478$	$+282$	$(+140000)$
	4^{n+1}	$+2.67$	$+4.61$	$+7.48$	$+1.10$	$(+137)$
	$4^{n+1}C_F C_A^n$	$\mathbf{+2.00}$	$\mathbf{+1.15}$	$\mathbf{+0.62}$	$\mathbf{+0.03}$	$\mathbf{(+1.27)}$

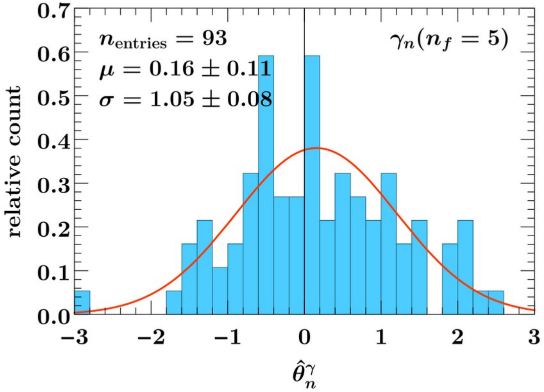
⋮

"Statistic of many computations"

Matrix elements



Anomalous dimensions



TNP approach for fixed-order computations

[Lim, Poncelet, 2412.14910]

$$\begin{aligned}
\mathrm{d}\sigma &= \alpha_s^n N_c^m \,\mathrm{d}\bar{\sigma}^{(0)} + \alpha_s^{n+1} N_c^{m+1}\,\mathrm{d}\bar{\sigma}^{(1)} + \alpha_s^{n+2} N_c^{m+2}\,\mathrm{d}\bar{\sigma}^{(2)} + \dots \\
&= \alpha_s^n N_c^m \,\mathrm{d}\bar{\sigma}^{(0)} \left[1 + \alpha_s N_c \left(\frac{\mathrm{d}\bar{\sigma}^{(1)}}{\mathrm{d}\bar{\sigma}^{(0)}}\right) + \alpha_s^2 N_c^2 \left(\frac{\mathrm{d}\bar{\sigma}^{(2)}}{\mathrm{d}\bar{\sigma}^{(0)}}\right) + \dots\right]
\end{aligned}$$

Observation, i.e. “expert knowledge”: $\dfrac{\mathrm{d}\bar{\sigma}^{(i)}}{\mathrm{d}\bar{\sigma}^{(0)}} \sim \mathcal{O}(1)$

Use some knowledge about lower orders but introduce parametric dependence:

$$\frac{\mathrm{d}\bar{\sigma}^{(N+1)}_{\mathrm{TNP}}}{\mathrm{d}\bar{\sigma}^{(0)}} = \sum_{j=1}^{N} f_k^{(j)}\left(\vec{\theta}, x\right)\left(\frac{\mathrm{d}\bar{\sigma}^{(j)}}{\mathrm{d}\bar{\sigma}^{(0)}}\right)$$

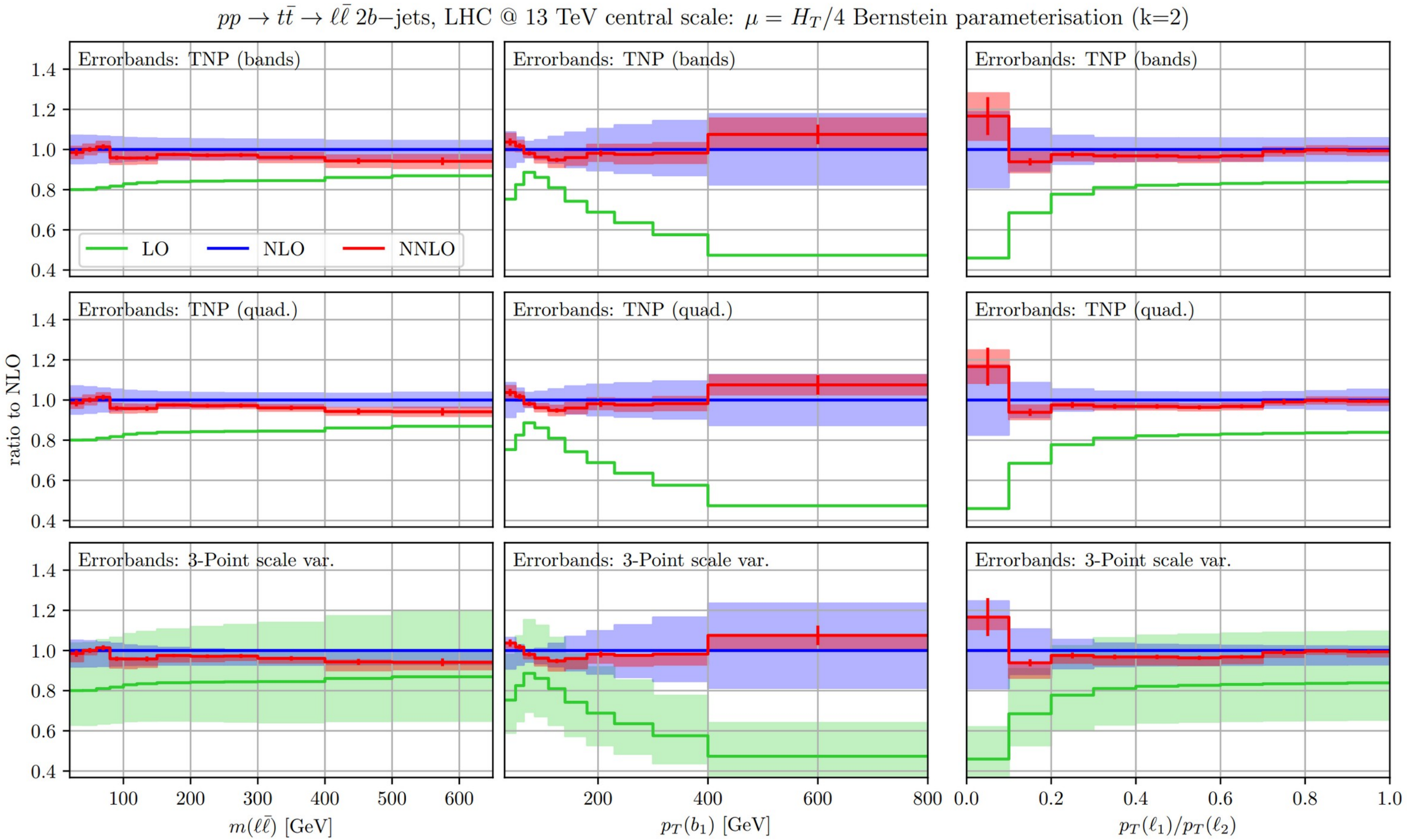
x → mapped kinematic variable

Approximation of original TNP philosophy
→ there is only $f_i(\hat{\theta}) \approx \hat{f}_i$

Bernstein: $f_k^B\left(\vec{\theta}, x\right) = \sum_{i=0}^{k} \theta_i \binom{k}{i} x^{k-i}(1-x)^i$, $x \in [0,1]$

Chebyshev: $f_k^C\left(\vec{\theta}, x\right) = \frac{1}{2}\sum_{i=0}^{k} \theta_i T_i(x)$, $x \in [-1,1]$

Uncertainties from TNPs - ttbar+decays



Band: sample

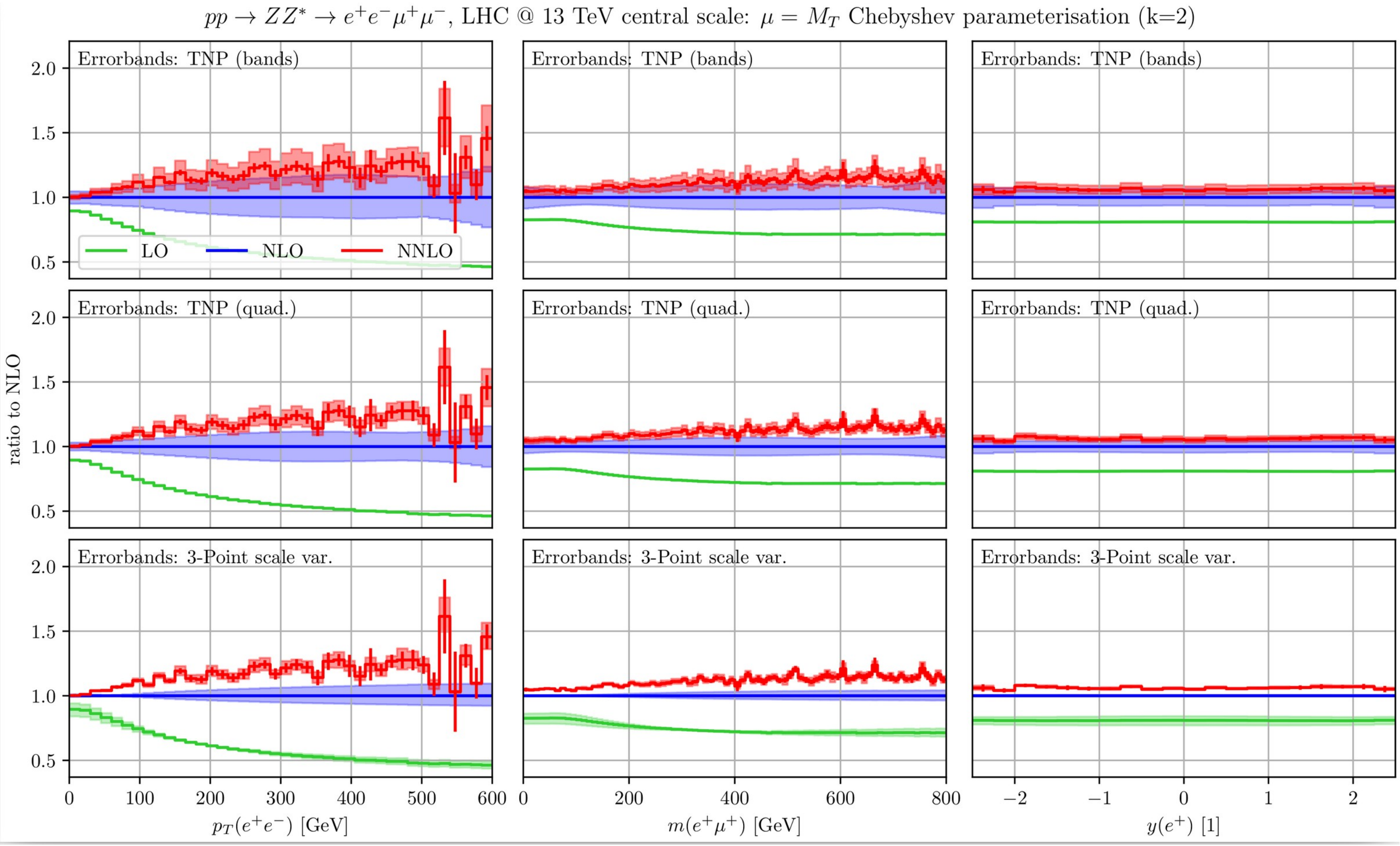
$\theta \in [-1, 1]$

Quad: add individual

$\theta = \pm 1$

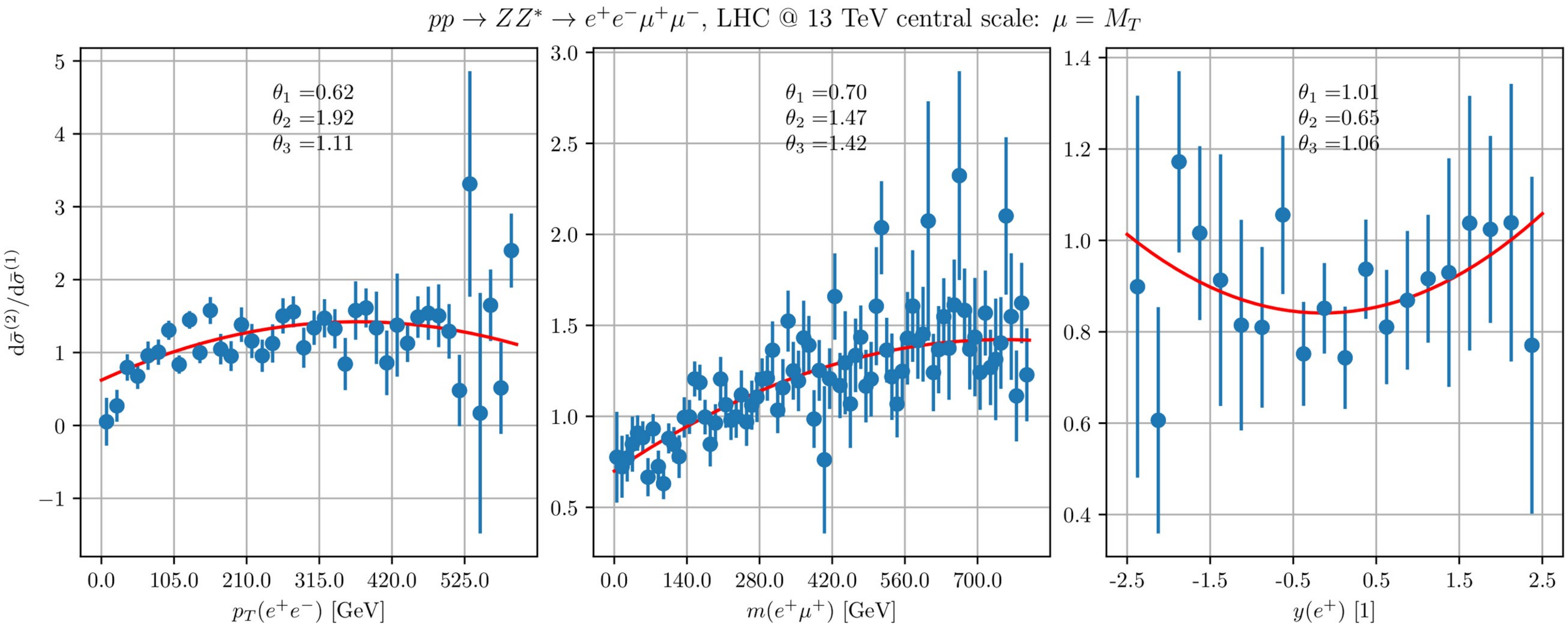
in quadrature

Uncertainties from TNPs - ZZ



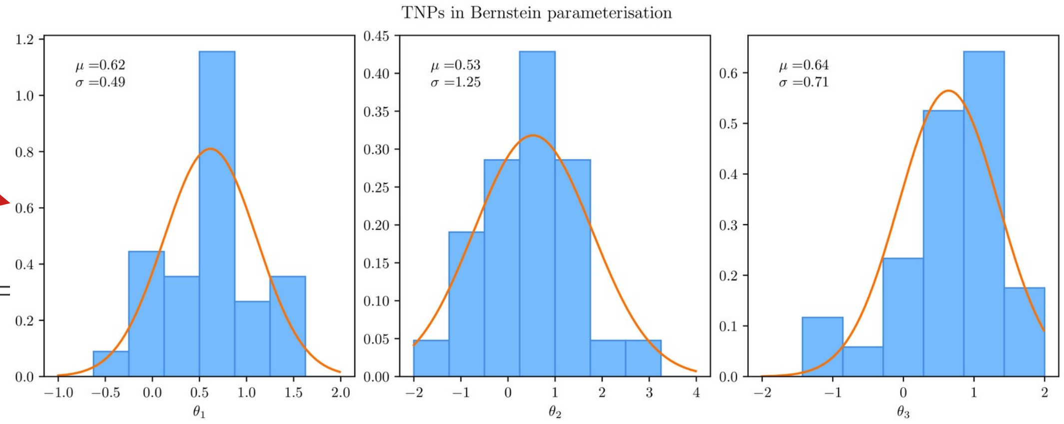
Example of TNP fit: pp → ZZ

$$\frac{d\bar{\sigma}^{(N+1)}_{\text{TNP}}}{d\bar{\sigma}^{(0)}} = \sum_{j=1}^{N} f_k^{(j)}\left(\vec{\theta}, x\right)\left(\frac{d\bar{\sigma}^{(j)}}{d\bar{\sigma}^{(0)}}\right) \qquad f_k^B\left(\vec{\theta}, x\right) = \sum_{i=0}^{k} \theta_i \binom{k}{i} x^{k-i}(1-x)^i$$



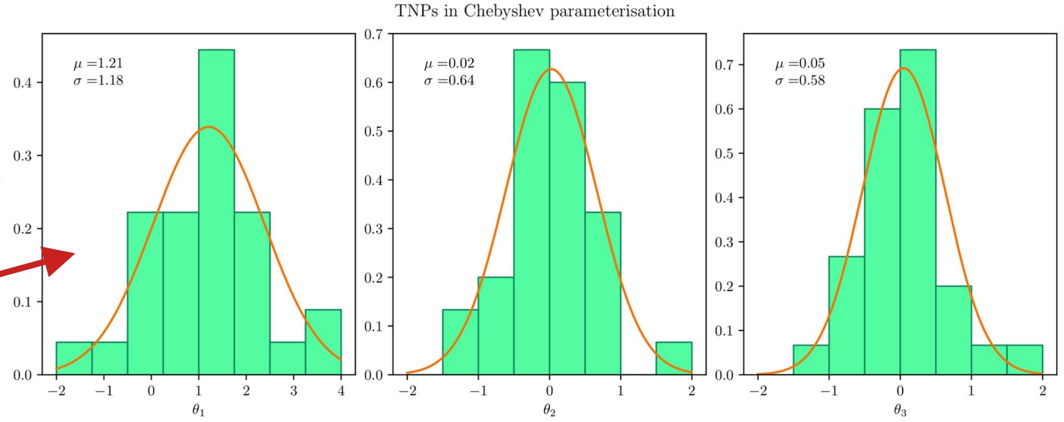
Fits - meta-study

$$f_k^B\left(\vec{\theta},x\right)=\sum_{i=0}^{k}\theta_i\begin{pmatrix}k\\ i\end{pmatrix}x^{k-i}(1-x)^i$$



Process	Distributions
$pp \to H$ (full theory)	y_H
$pp \to ZZ^* \to e^+e^-\mu^+\mu^-$	$M_{e^+\mu^+}$, $p_T^{e^+e^-}$, y_{e^+}
$pp \to WW^* \to e\nu_e\mu\nu_\mu$	M_{WW}, $p_T^{\mu^-}$, y_{W^-}
$pp \to (W \to \ell\nu) + c$	p_T^ℓ ,$\lvert y_\ell\rvert$,
$pp \to t\bar{t}$	$M_{t\bar{t}}$, p_T^t, y_t
$pp \to t\bar{t} \to b\bar{b}\ell\bar{\ell}$	$M_{\ell\bar{\ell}}$, $p_T^{b_1}$, $p_{T,\ell_1}/p_{T,\ell_2}$
$pp \to \gamma\gamma$	$M_{\gamma\gamma}$, $p_T^{\gamma_1}$, $y_{\gamma\gamma}$
$pp \to \gamma\gamma j$	$M_{\gamma\gamma}$, $p_T^{\gamma\gamma}$, $\cos\phi_{\mathrm{CS}}$, $\lvert y_{\gamma_1}\rvert$, $\Delta\phi_{\gamma\gamma}$
$pp \to jjj$	EEC with $H_{T,2} \in [1000, 1500), [1500, 2000), [3500, \infty)$
$pp \to \gamma jj$	$M_{\gamma jj}$, p_T^j, $\lvert y_{\gamma-\mathrm{jet}}\rvert$, $E_{T,\gamma}$

$$f_k^C\left(\vec{\theta},x\right)=\frac{1}{2}\sum_{i=0}^{k}\theta_i T_i(x)$$



Points for discussion, caveats and open questions

Some arising questions regarding fixed-order model: $$\frac{\mathrm{d}\bar{\sigma}^{(N+1)}_{\mathrm{TNP}}}{\mathrm{d}\bar{\sigma}^{(0)}} = \sum_{j=1}^{N} f_k^{(j)}\left(\vec{\theta}, x\right)\left(\frac{\mathrm{d}\bar{\sigma}^{(j)}}{\mathrm{d}\bar{\sigma}^{(0)}}\right)$$

- How does the uncertainty estimate depend on the central scale choice?
  → **bad scale choices lead to large uncertainties by construction due to large corrections.**
- What about NLO uncertainty if $\mathrm{d}\bar{\sigma}^{(1)} = 0$ for given scale?
  → **amend parametrisation by j = 0 term.**
- Each parametrisation is for one observable at a time:
  How to deal with higher dimensional distributions? Consistency upon integration?
  → **WIP**
- What about EW corrections?
  → **Sudakov logs should work well!** → **Radiation from resonances more difficult.**
- How to correlate different processes at fixed-order?
  → **???, would require something like:** $\mathrm{d}\bar{\sigma}^{(n)}(\theta) = \mathrm{d}\Phi \left\langle M^0 \middle| \mathcal{P}(\theta) \middle| M^0 \right\rangle$ $\mathcal{P}(\theta)$ → process-independent "operator"
- How sensitive are we to the parametrisation? How many terms?
  → **two quite general parametrisations tested, increase degree by demand.**

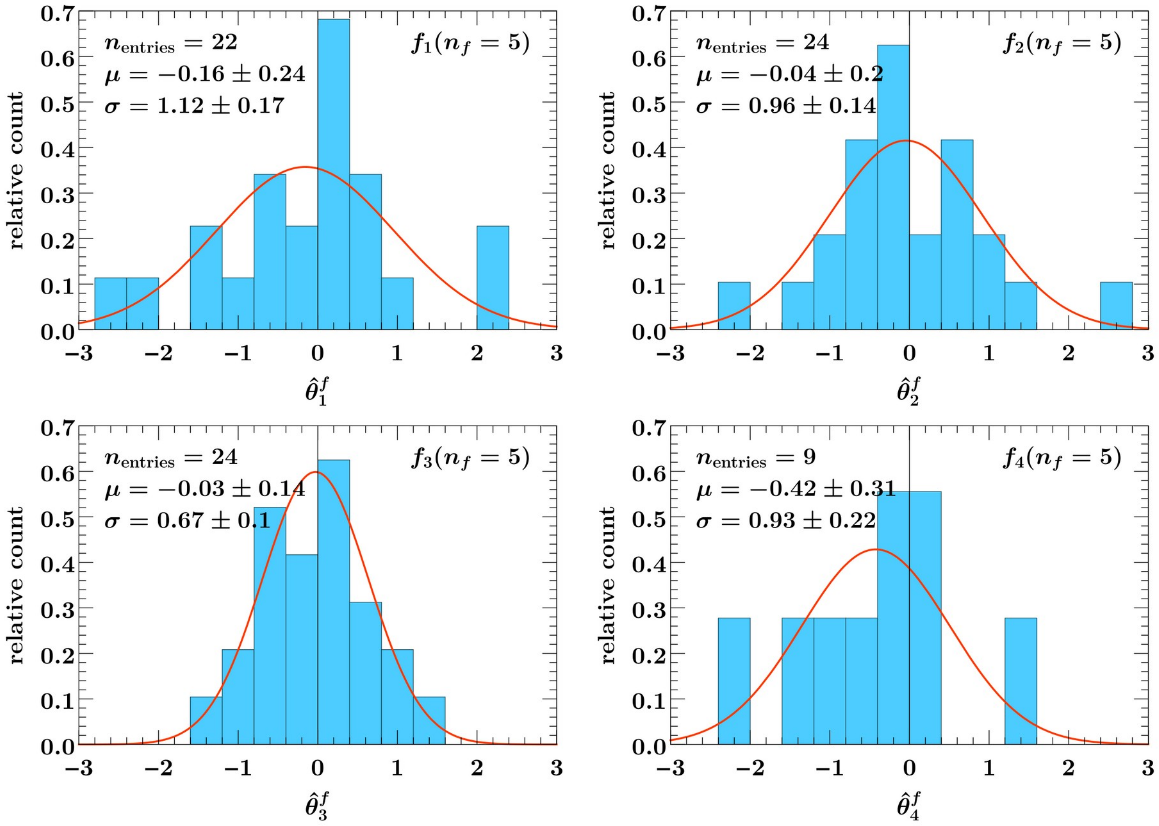
Take home message

- Increasing precision demands accurate theory uncertainty estimates
- De-facto standard: scale variations
  → various short-comings: robustness, no statistical interpretation, correlations,...
- Alternative approaches to scale variations: Bayesian and TNP approach
- **Theory Nuisance Parameters**
  - In principle less biased, better correlations → does not depend on any "known" orders
    … however needs "expert knowledge"
  - Allows for a statistical interpretation and constraints from data!
  - Fixed-order tricky, not much knowledge about higher-order terms
    Proposed TNP parametrisation of differential cross sections shows promising first results
    next step: application to an actual parameter fit

Is this the ultimate answer? Surely not, but a step in the right direction!

Backup

TNP parametrisations for resummation



Higgs pT spectrum

[Cal,Lim, Scott,Tackmann Waalewijn 2408.13301]

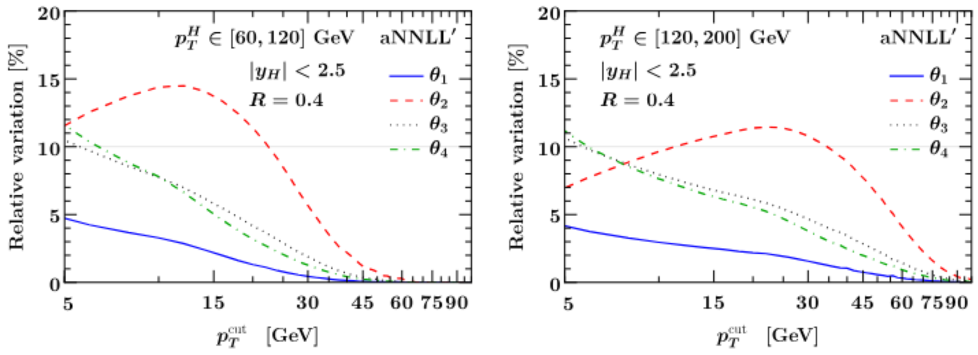
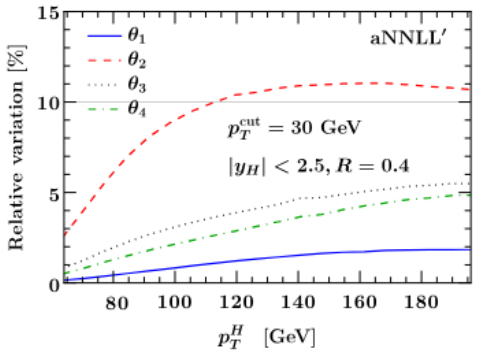


Figure 6: Relative uncertainty from varying each theory nuisance parameter as a function of p_T^{cut} for two different STXS bins.



Example: incomplete knowledge of NNLL resummation

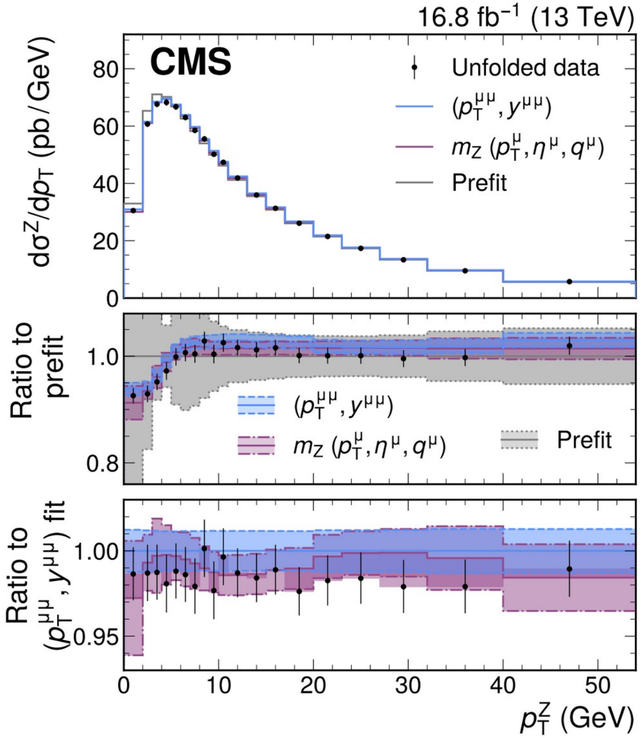
→ some two-loop ingredients unknown
→ parametrise by TNPs

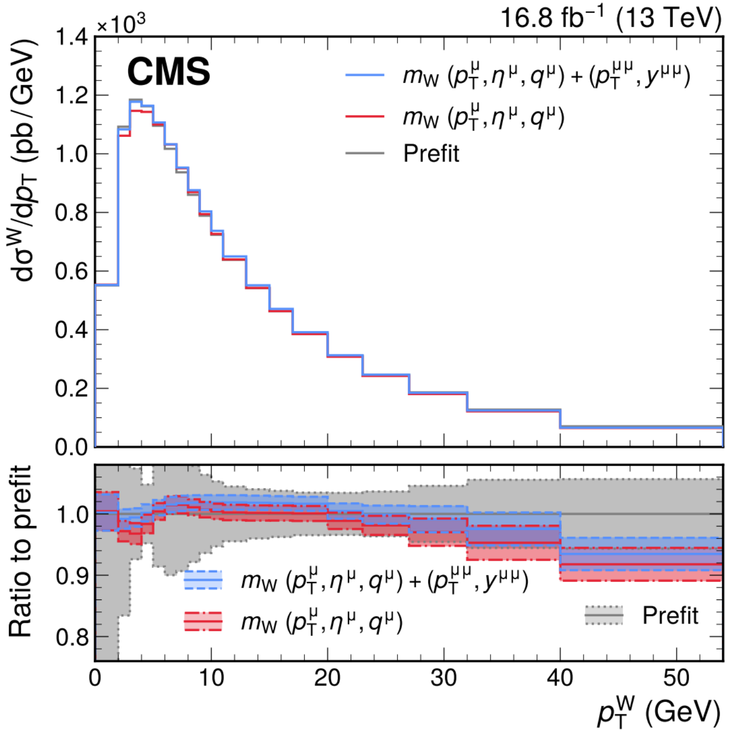
→ Make predictions and vary TNPs:

- correlated uncertainty for different bins!
- See impact of different missing ingredients

Constraint of TNPs from data → W-mass extraction

Resummation ingredients the same for W or Z production
→ constraint from precisely measured pT(Z) → use for pT(W)
→ massive reduction of unc. with correct correlations!





Some remarks on TNPs in resummation

Picture: simple ingredients that enter different computations/processes etc.

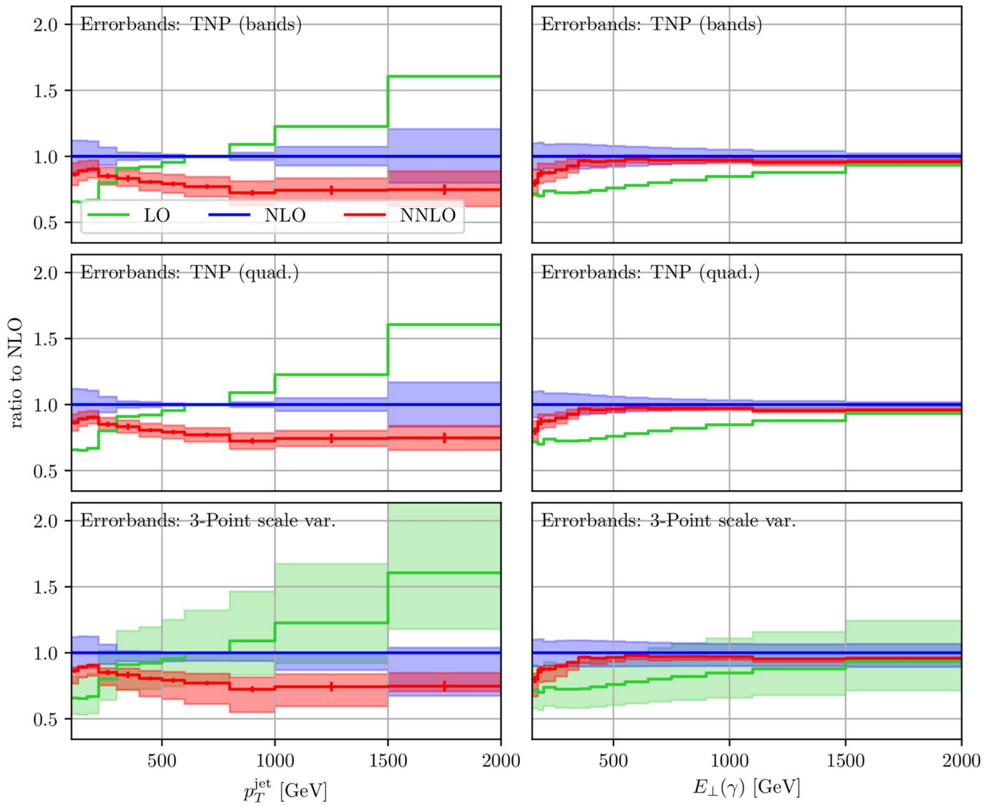
→ ideal situation

But actually not that simple:

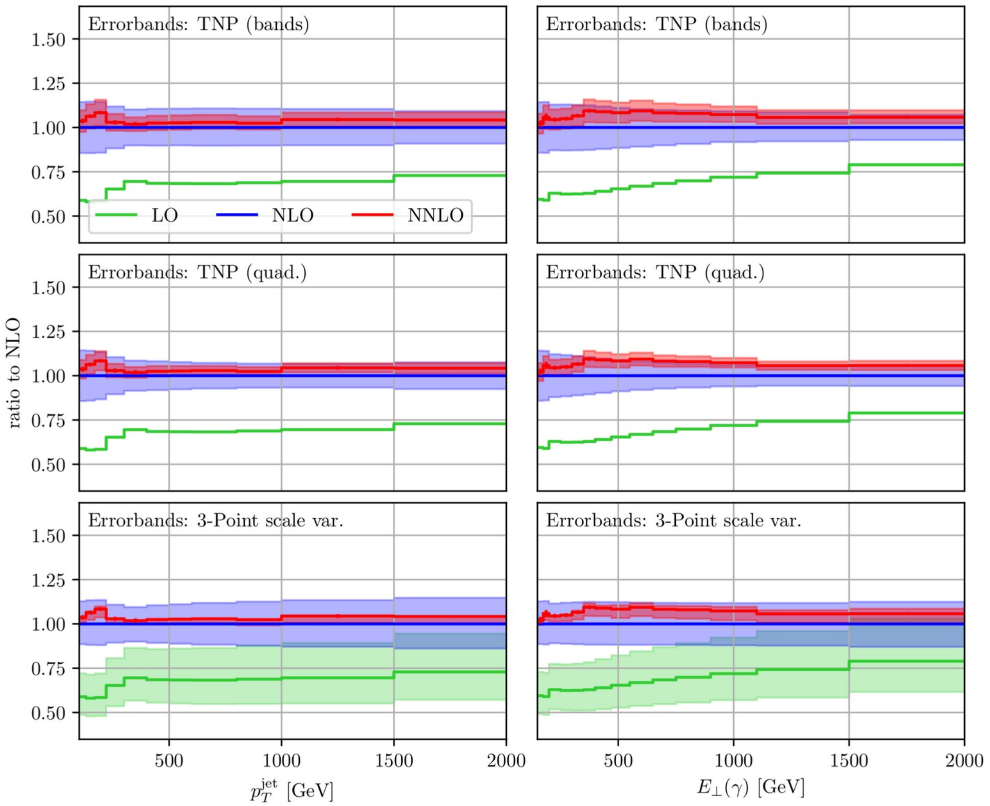
- Scheme dependence of ingredients? E.g. scales, IR subtraction, …
  → might need modified parametrisations
- Some TNPs represent directly numbers: $\Gamma,\ \gamma,\ H$ for simple processes
  but others are functions → Beam functions, hard functions for more complicated processes
- Parametrisation works for the resummed spectrum. What about other observables?
- “Easy to implement” for use-cases so far
  → might be really expensive if each variation needs a full computations (Monte Carlos,...)

Challenging case

“Bad” scale choice $\mu = E_T$ no j=0 term

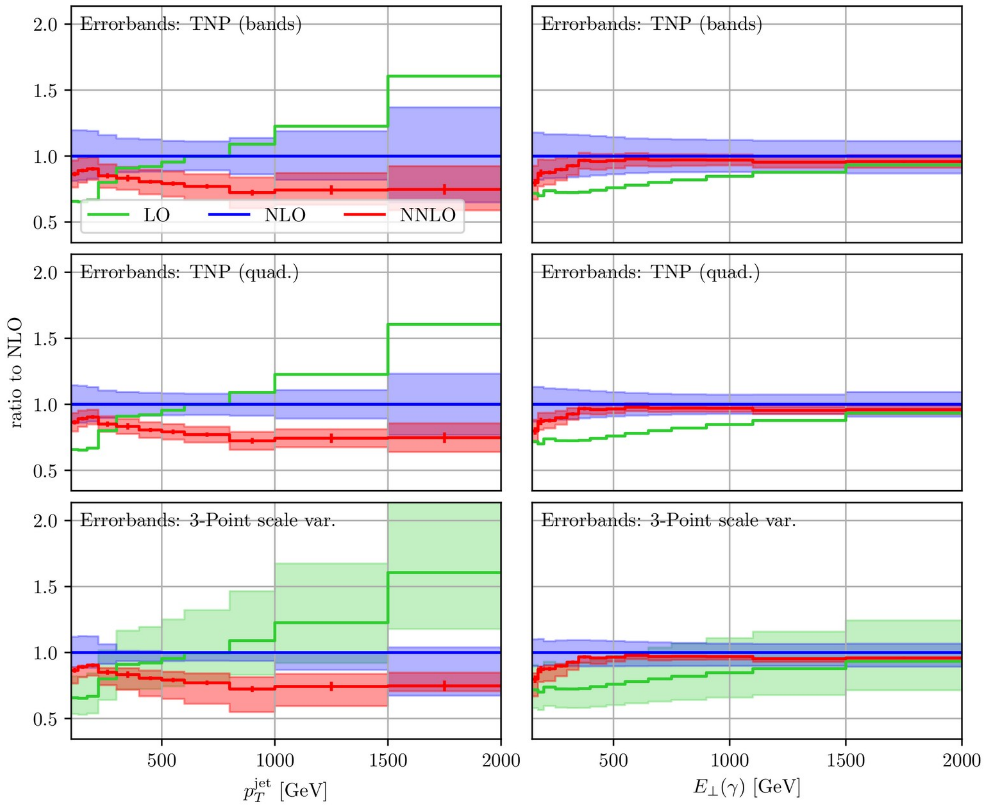


“Good” choice $\mu = H_T$ no j=0 term

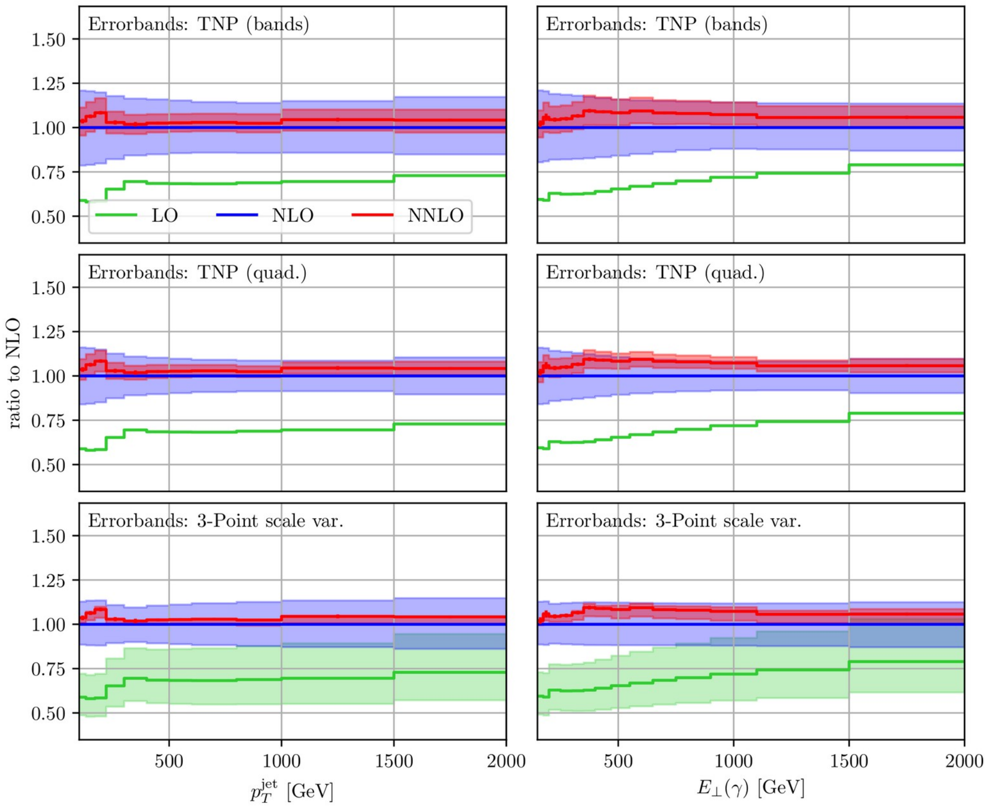


Challenging case → extended parametrisations

“Bad” scale choice $\mu = E_T$ with j=0 term

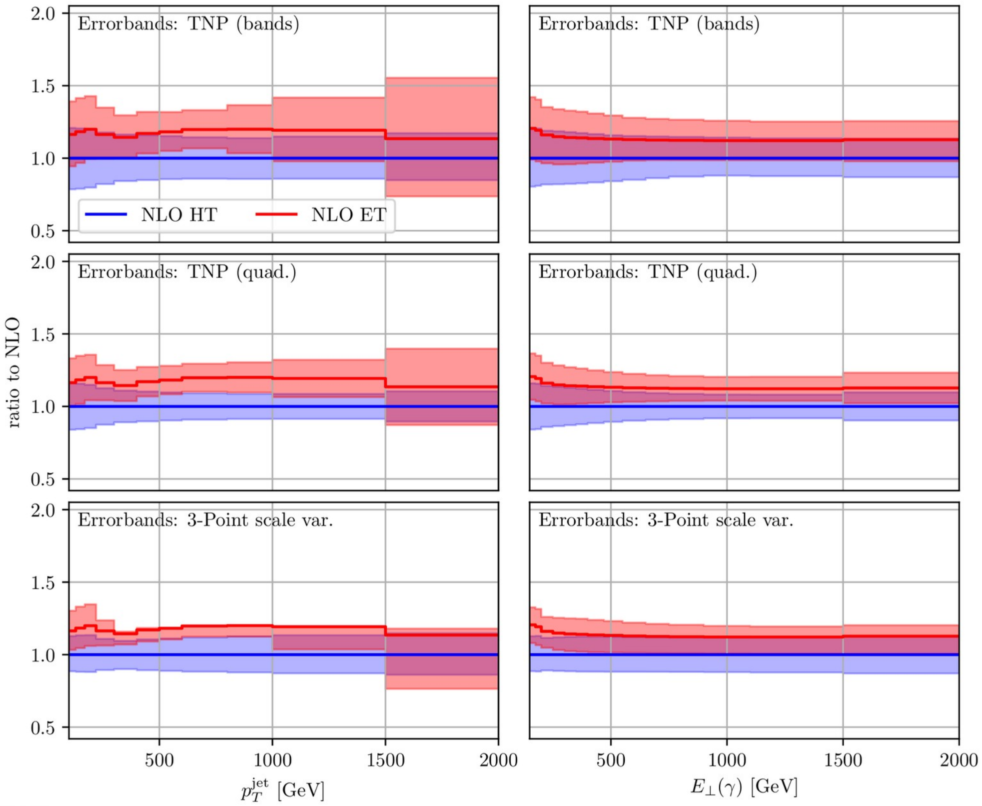


“Good” choice $\mu = H_T$ with j=0 term

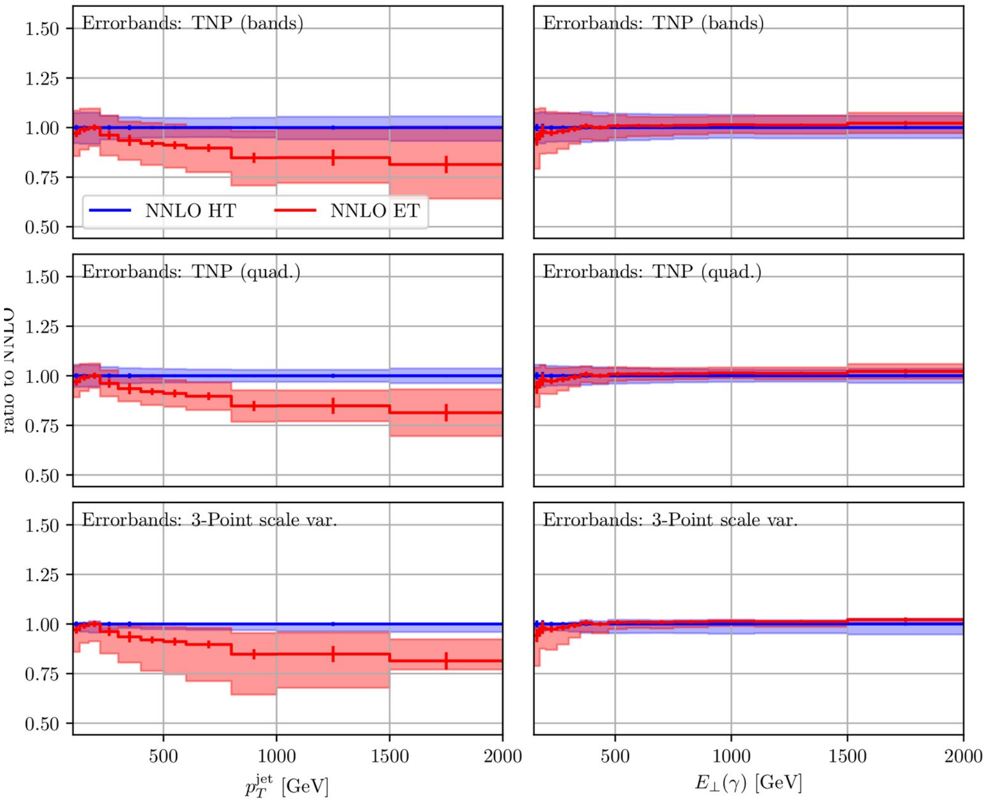


Challenging case → comparisons

NLO QCD



NNLO QCD



Bayesian approach I

→ Instead of ad-hoc fixed variation try to give some probabilistic interpretation

$$\mathrm{d}\sigma = \mathrm{d}\sigma^{(0)}(1 + \delta^{(1)} + \delta^{(2)} + \dots)$$

Probability to find coefficient $\delta^{(n+1)}$ given $\delta^{(n)}$:

[Cacciari,Houdeau 1105.5152]

$$P(\delta^{(n+1)}|\delta^{(n)}) = \frac{P(\delta^{(n+1)})}{P(\delta^{(n)})} = \frac{\int \mathrm{d}a P(\delta^{(n+1)}|a) P_0(a)}{\int \mathrm{d}a P(\delta^{(n)}|a) P_0(a)}$$

Need to provide model and prior

Bayes: $P(A|B) = P(B|A)P(A)/P(B)$ with: $P(\delta^{(n)}|\delta^{(n+1)}) = 1$

CH model: $\delta_k = c_k \alpha_s^k$ c_k come from geometric series: $|c_k| \leq \bar{c} \quad \forall k$

Geometric model: $|\delta_k| \leq c a^k \quad \forall k$ [Bonvini 2006.16293]

abc model: $b - c \leq \frac{\delta_k}{a^k} \leq b + c \quad \forall k$ [Duhr, Huss, Mazeliauskas, Szafron 2106.04585]

Bayesian approach II

Inclusion of scale dependence: $$P(\delta_{n+1}|\delta_n) = \int \mathrm{d}\mu P(\delta_{n+1}|\delta_n;\mu)P(\mu|\delta_n)$$

Scale marginalisation (the scale becomes a model parameter)

$$\mathcal{P}_{\mathrm{sm}}(\Sigma|\boldsymbol{\Sigma}_n) \approx \frac{\int d\mu\, P(\Sigma - \Sigma_n(\mu)|\boldsymbol{\Sigma}_n(\mu))\, P(\boldsymbol{\Sigma}_n(\mu))\, P_0(\mu)}{\int d\mu'\, P(\boldsymbol{\Sigma}_n(\mu'))\, P_0(\mu')}\,.$$

μ_{FAC}

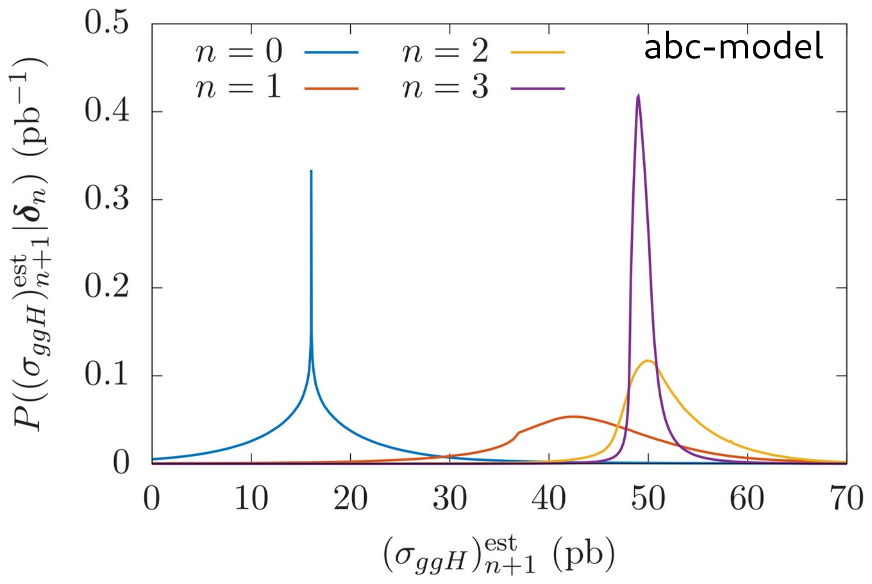
Scale average (the results are averaged with weight function)

$$\mathcal{P}_{\mathrm{sa}}(\Sigma|\boldsymbol{\Sigma}_n) \approx \int d\mu\, w(\mu)\, P(\Sigma - \Sigma_n(\mu)|\boldsymbol{\Sigma}_n(\mu))$$

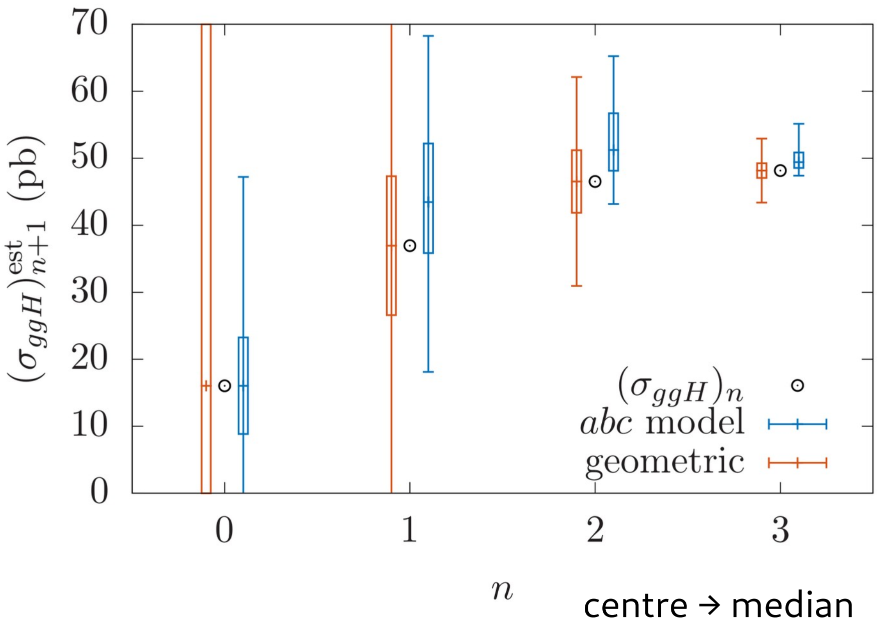
μ_{PMS}

Bayesian approach III

Example: Higgs production in gluon - fusion



Comparison of different unc. estimates:



centre → median
error box → 68% CI
error bar → 95% CI