

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

---

Rene Poncelet



THE HENRYK NIEWODNICZAŃSKI  
INSTITUTE OF NUCLEAR PHYSICS  
POLISH ACADEMY OF SCIENCES

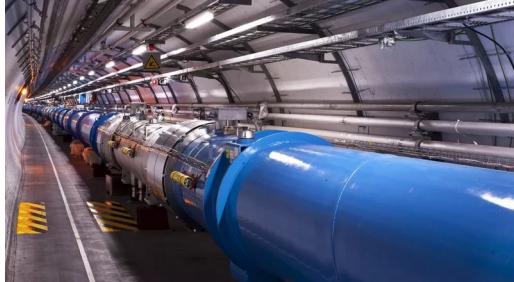
# Outline

---

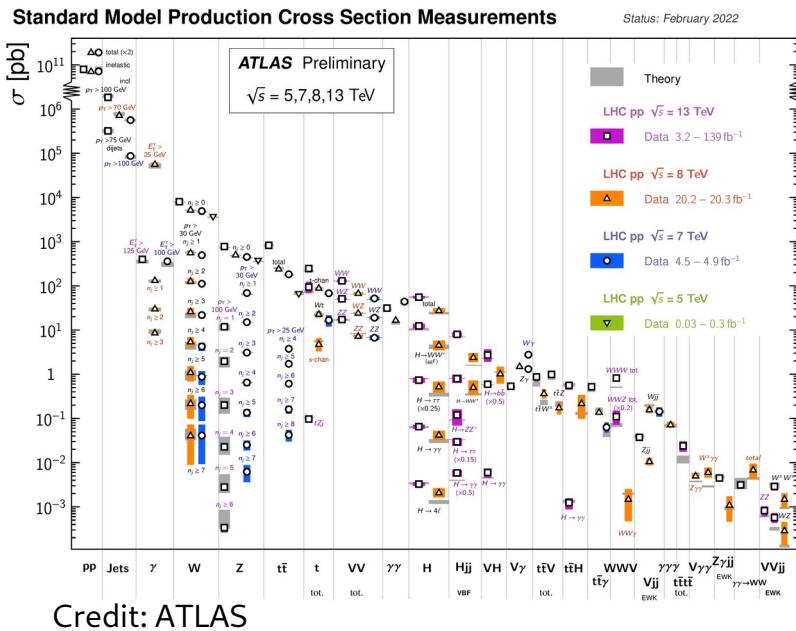
- Precision predictions at the LHC
- Missing Higher Order Uncertainties (MHOU)  
How to estimate the uncertainty of (truncated) perturbative expansions?
  - Scale variations for fixed-order and resummed cross sections
  - Bayesian methods
  - Theory Nuisance Parameters (TNPs)
- Application of TNPs to fixed-order perturbation theory
- Discussion/Summary/Outlook

# Standard Model phenomenology at the LHC

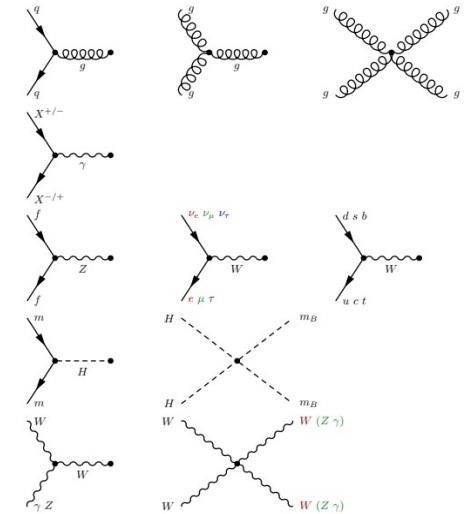
Scattering experiments



Credit: CERN



Theory/Model

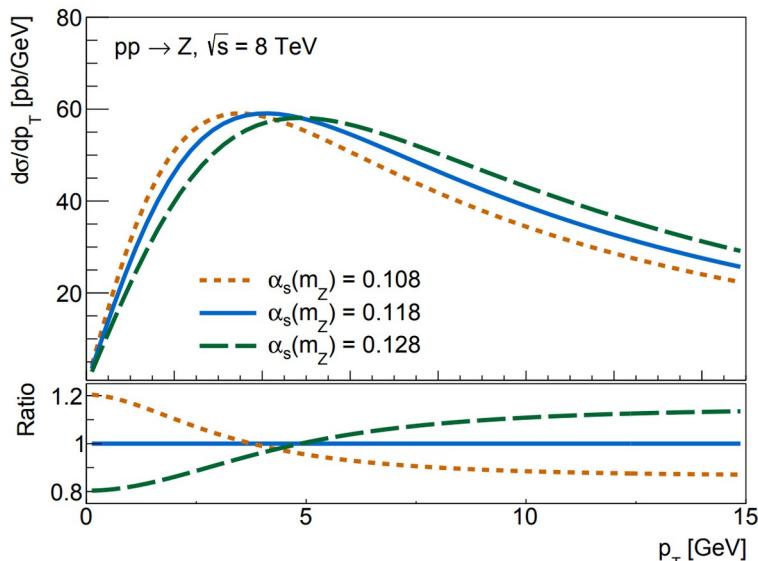


Credit: Jack Lindon, CERN

# Precision example: strong coupling from pT(Z)

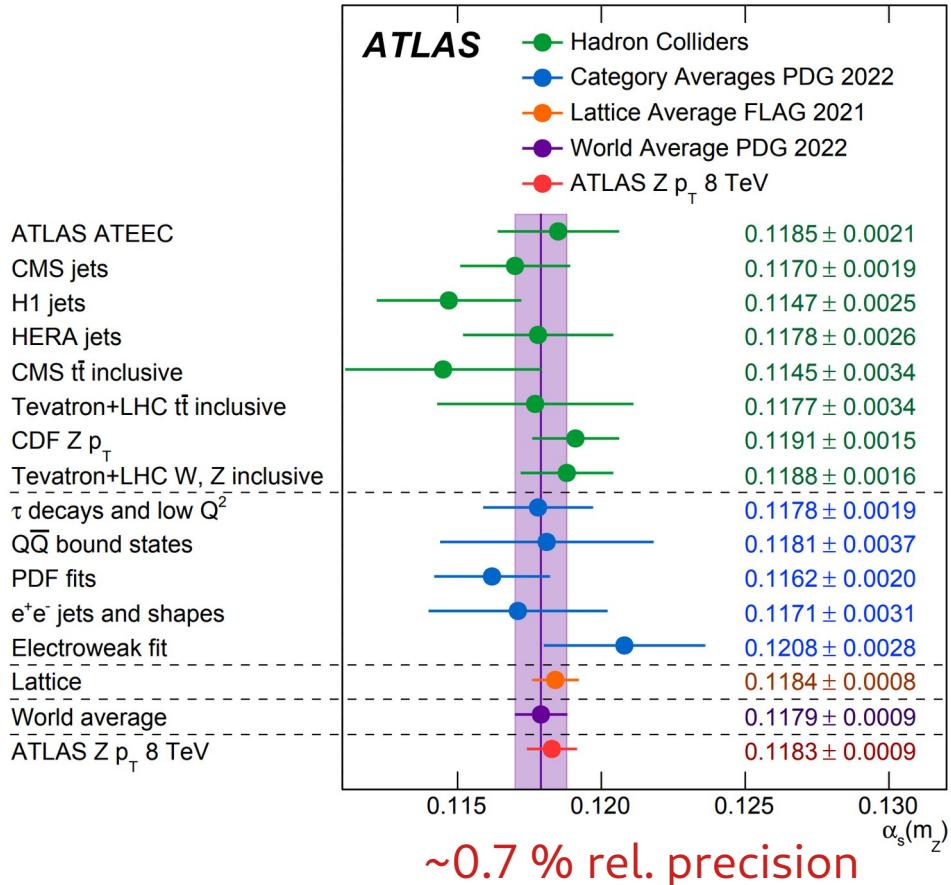
[ATLAS 2309.12986]

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



→ at low pT resummation regime!

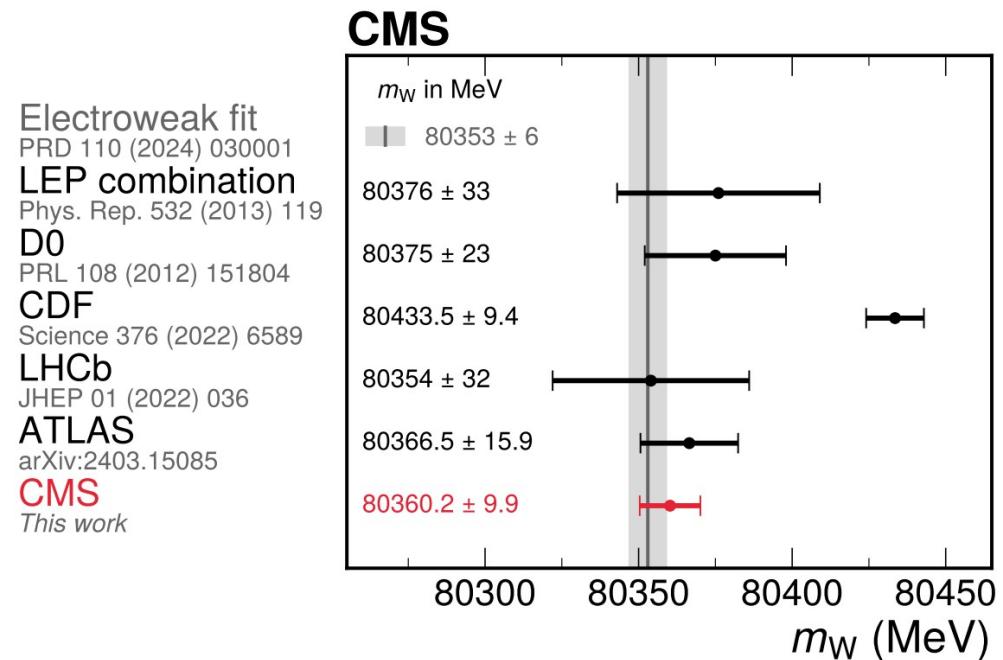
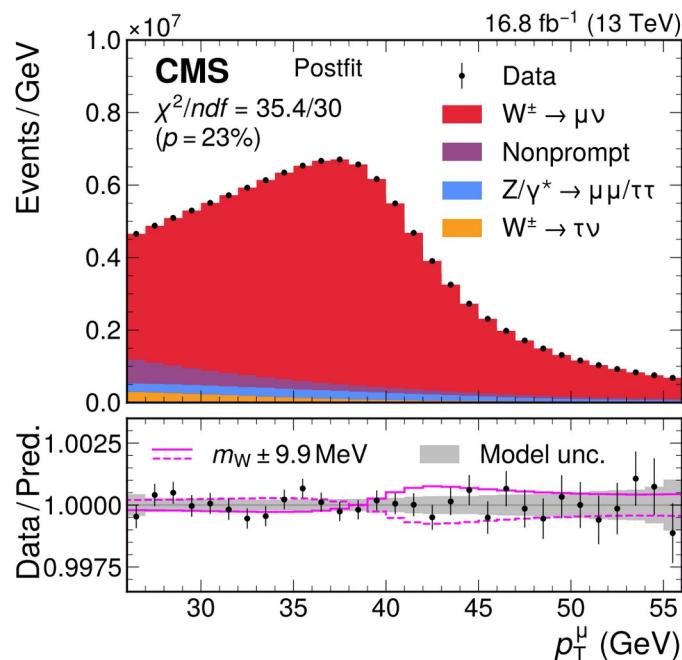
→ theory uncertainty?



# Precision example: W-mass measurement by CMS

[CMS 2412.13872]

Mass dependence of  $p_T(l)$ :

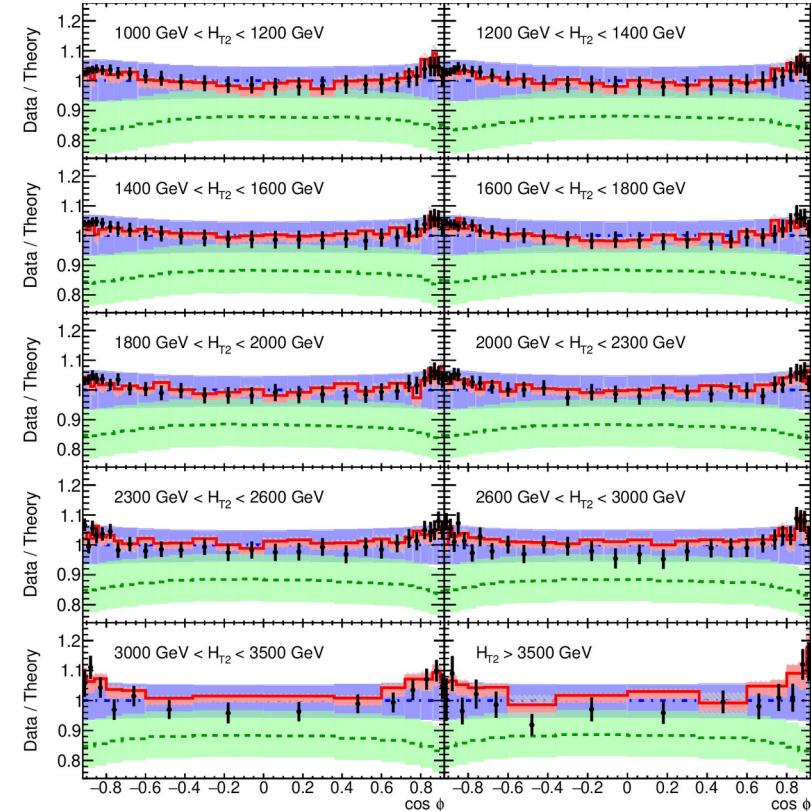
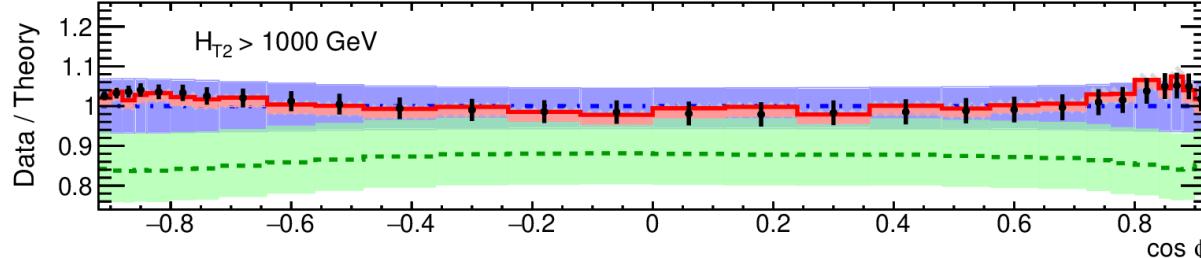


Jacobian peak position  $\sim m(W)/2 \rightarrow$  resummation sensitive  $\rightarrow$  theory uncertainty?

# Precision example: strong-coupling from TEEC

Next-to-Next-to-Leading Order Study of Three-Jet Production at the LHC  
 Czakon, Mitov, Poncelet Phys.Rev.Lett. 127 (2021) 15, 152001

[ATLAS 2301.09351]



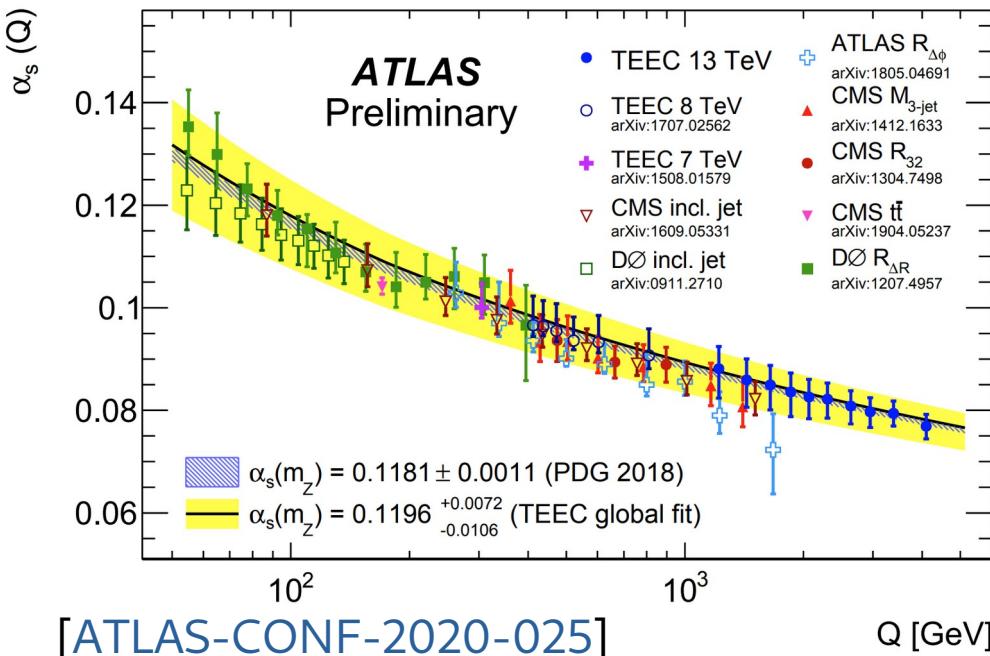
Multi-jet angular correlations

Uncertainties driven by  
 fixed-order precision through ratio:

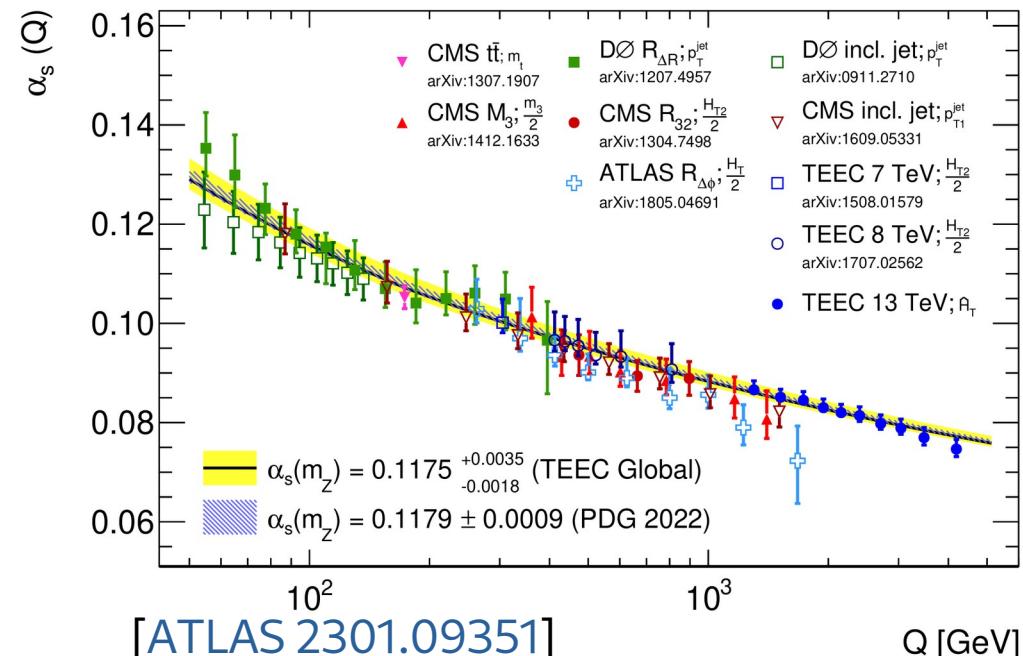
$$R^i(\mu_R, \mu_F, \text{PDF}, \alpha_{S,0}) = \frac{d\sigma_3^i(\mu_R, \mu_F, \text{PDF}, \alpha_{S,0})}{d\sigma_2^i(\mu_R, \mu_F, \text{PDF}, \alpha_{S,0})}$$

# Precision example: strong-coupling from TEEC

NLO QCD

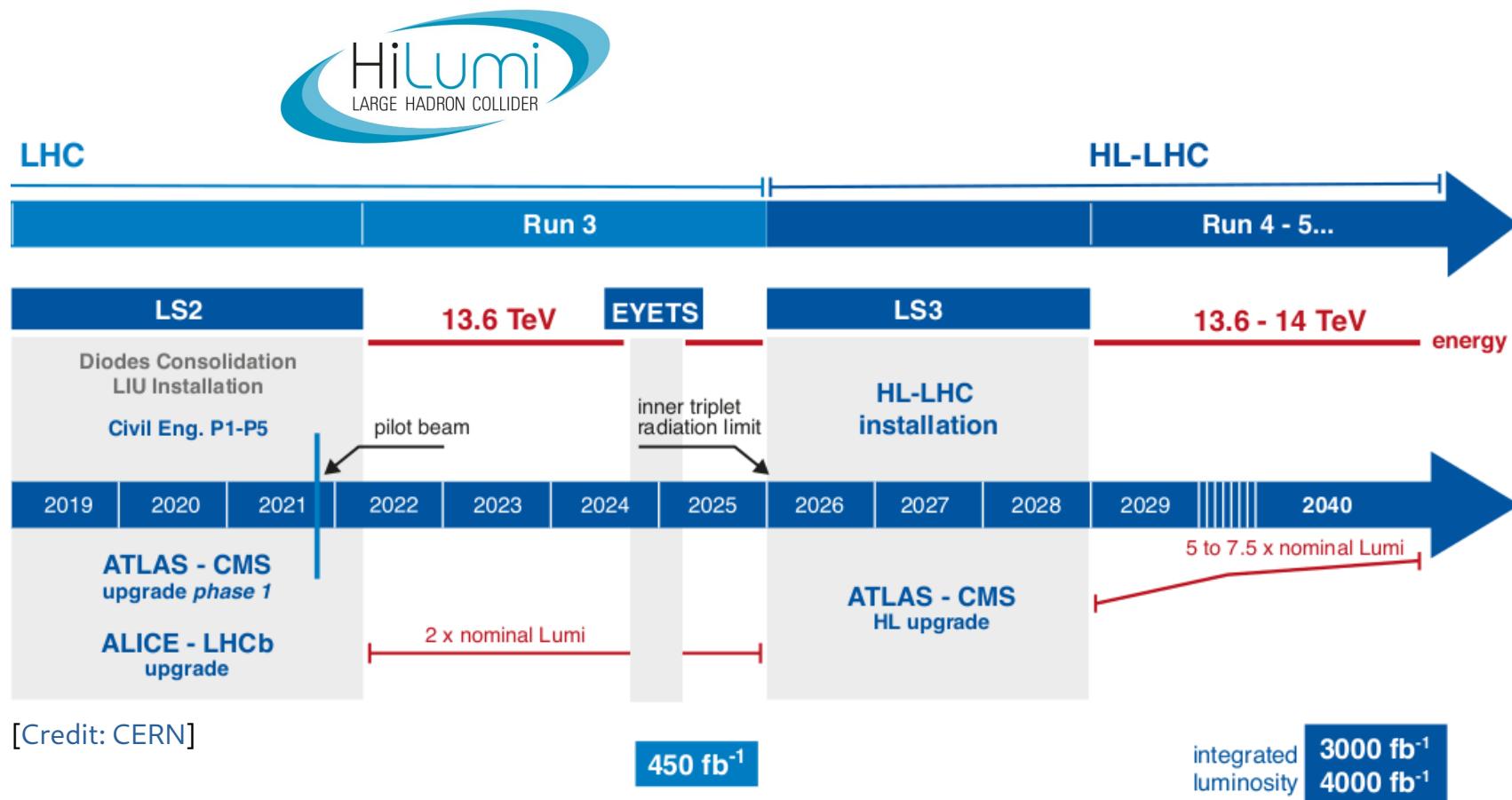


NNLO QCD



Theory uncertainty dominant effect!

# LHC Precision era and future experiments

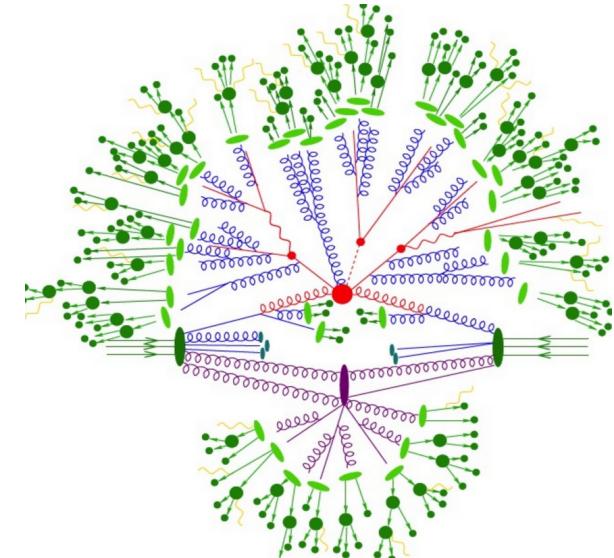


# Uncertainties in precision phenomenology

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]

# Theory uncertainties from scale variations

Lets focus on QCD as an example:

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

RGE

$$\mu \frac{d\sigma^{(n)}}{d\mu} = \mathcal{O}\left(\alpha_s^{(n_0+n+1)}\right)$$

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

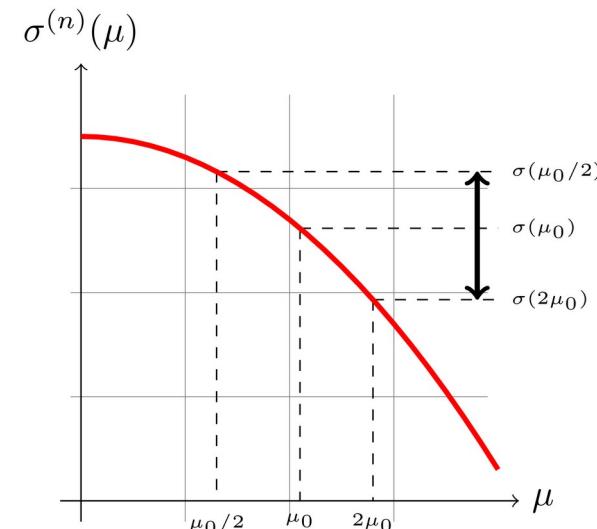
## Scale variation prescription

- choose 'sensible'  $\mu_0$

$\sigma^{(n)}(\mu_{\text{FAC}}) = 0 \rightarrow$  principle of fastest apparent convergence

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

- vary with a factor (typically 2)
- take envelope as uncertainty
- ad-hoc and heuristic choice



# Missing higher orders

---

Generic perturbative expansion:

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

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

For QCD:  $\alpha = \alpha_s(\mu_0)$

$f_i$  : the coefficient of the series, potentially unknown

We can compute the truncated series:

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

$\hat{f}_i$  : the true value, i.e. a value we actually computed

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

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

# Scale variation approach

---

Change of scale = change of renormalisation scheme:  $\tilde{\alpha}(\alpha) = \alpha(1 + b_0\alpha + b_1\alpha^2 + b_2\alpha^3 + \dots)$

For QCD:  $\alpha = \alpha_s(\mu_0)$      $\tilde{\alpha} = \alpha_s(\mu)$

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

$$\tilde{f}^{\text{LO}}(\tilde{\alpha}) = \hat{\tilde{f}}_0$$

$$\tilde{f}^{\text{NLO}}(\tilde{\alpha}) = \hat{\tilde{f}}_0 + \hat{\tilde{f}}_1 \tilde{\alpha} = \hat{f}_0 + \alpha \hat{f}_1 + \boxed{\alpha^2 b_0 \hat{f}_1} + \mathcal{O}(\alpha^3)$$

$$\tilde{f}^{\text{NNLO}}(\tilde{\alpha}) = \hat{\tilde{f}}_0 + \hat{\tilde{f}}_1 \tilde{\alpha} + \hat{\tilde{f}}_2 \tilde{\alpha}^2 = \hat{f}_0 + \alpha \hat{f}_1 + \alpha^2 \hat{f}_2 + \boxed{\alpha^3 (2b_0(\hat{f}_2 - b_0 \hat{f}_1) + b_1 \hat{f}_1)} + \mathcal{O}(\alpha^4)$$

# Scale variation approach

---

Change of scale = change of renormalisation scheme:  $\tilde{\alpha}(\alpha) = \alpha(1 + b_0\alpha + b_1\alpha^2 + b_2\alpha^3 + \dots)$

For QCD:  $\alpha = \alpha_s(\mu_0)$      $\tilde{\alpha} = \alpha_s(\mu)$

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

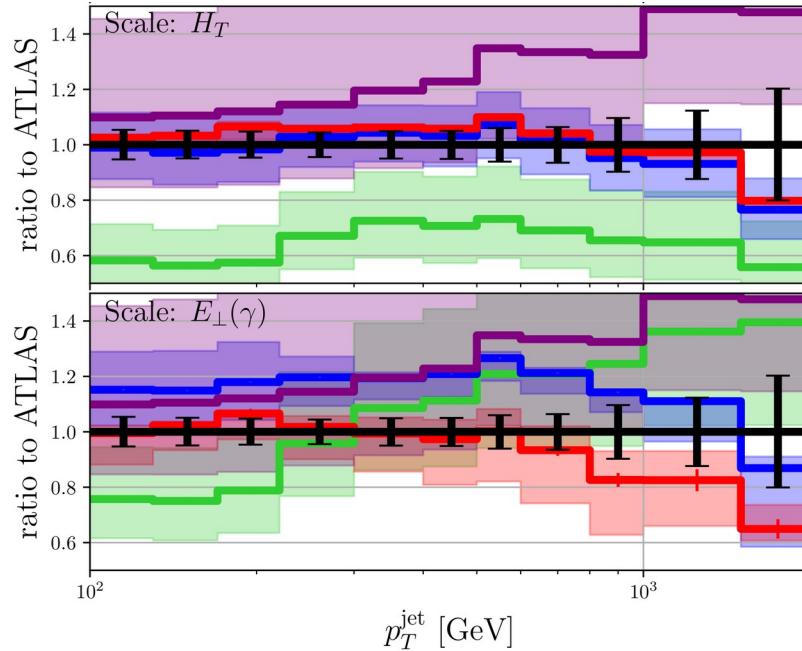
$$\Delta f^{\text{NLO}} = f^{\text{NLO}} - \tilde{f}^{\text{NLO}}(\tilde{\alpha}) = -\alpha^2 b_0 \hat{f}_1 + \mathcal{O}(\alpha^3)$$

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

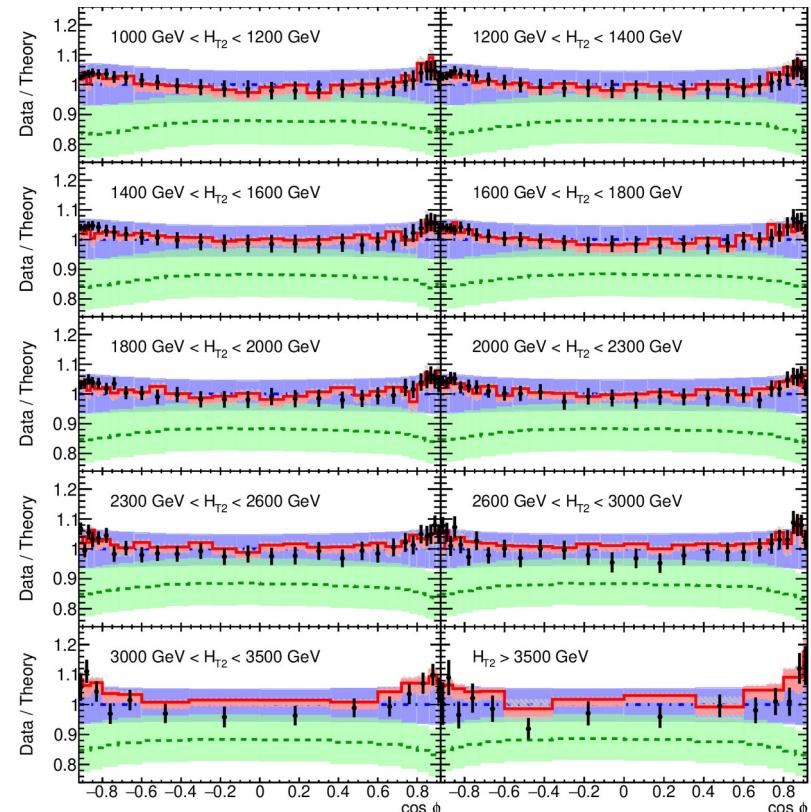
Issues:

- 1) There is no reason to believe that there is a value  $L$  (i.e. scale choice) that describes all  $\hat{f}_i$
- 2) If  $f$  is not a scalar, correlations are unclear

# Still, scale variation works ...

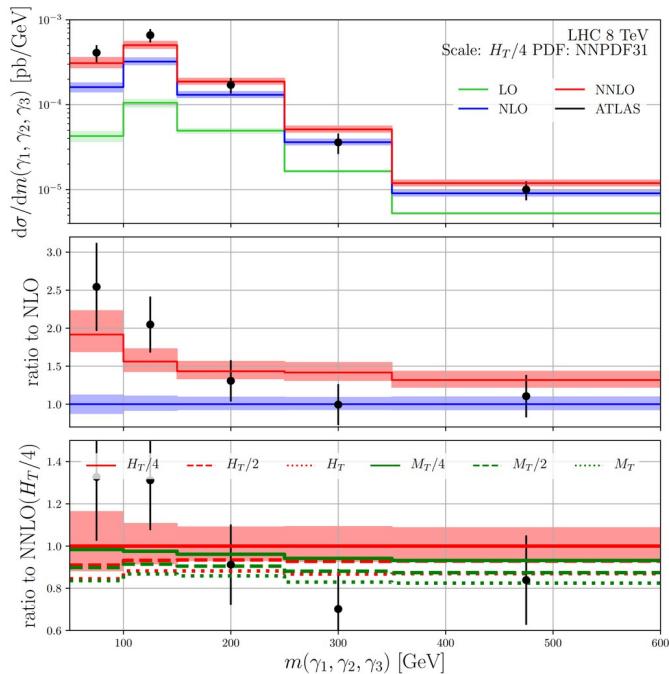


Badger, Czakon, Hartanto, Moodie, Peraro, **Poncelet**, Zoi  
[2304.06682 ]



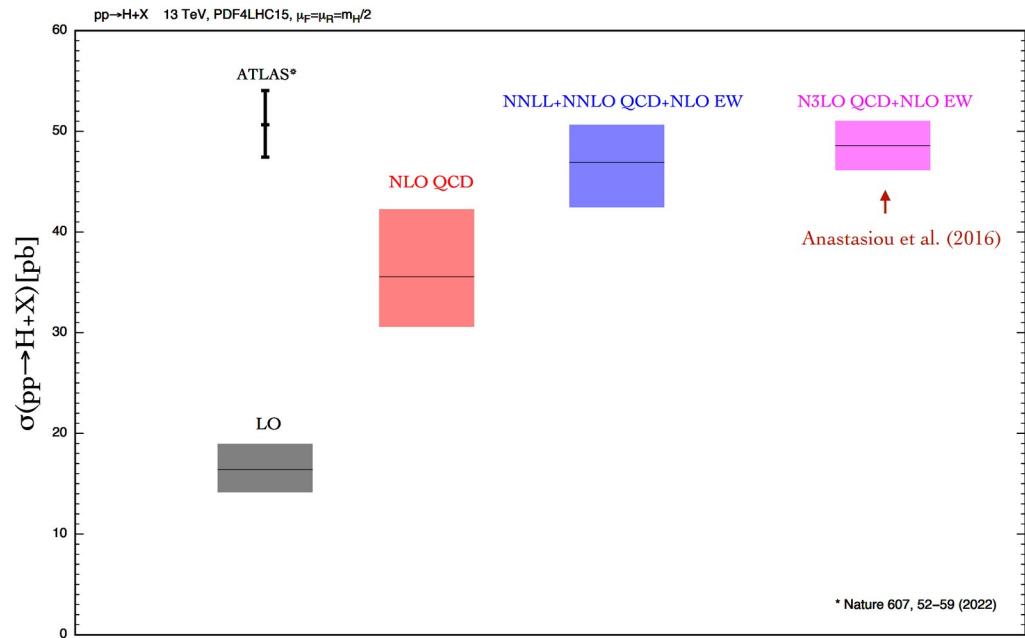
# ...sometimes :/

## Three photon production



NNLO QCD corrections to three-photon production at the LHC, Chawdhry, Czakon, Mitov, Poncelet  
[JHEP 02 (2020) 057]

## Higgs production



[talk by Grazzini]

# Short comings of scale variations

---

- not always reliable
- ...but in most cases issues are understood/expected:
  - new channels, phase space constraints, etc.
  - workarounds
- however, some issues are more fundamental:
  - how to choose the central scale? → not a physical parameter, no 'true' value  
(Principle of fastest 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” and “Theory Nuisance Parameter”

# Bayesian approach I

---

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

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

Probability to find coefficient  $\delta^{(n+1)}$  given  $\delta^{(n)}$ : [Cacciari,Houdeau 1105.5152]

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

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:  $|\delta_k| \leq ca^k \quad \forall k$  [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:

$$P(\delta_{n+1}|\delta_n) = \int d\mu P(\delta_{n+1}|\delta_n; \mu) P(\mu|\delta_n)$$

Scale marginalisation (the scale becomes a model parameter)

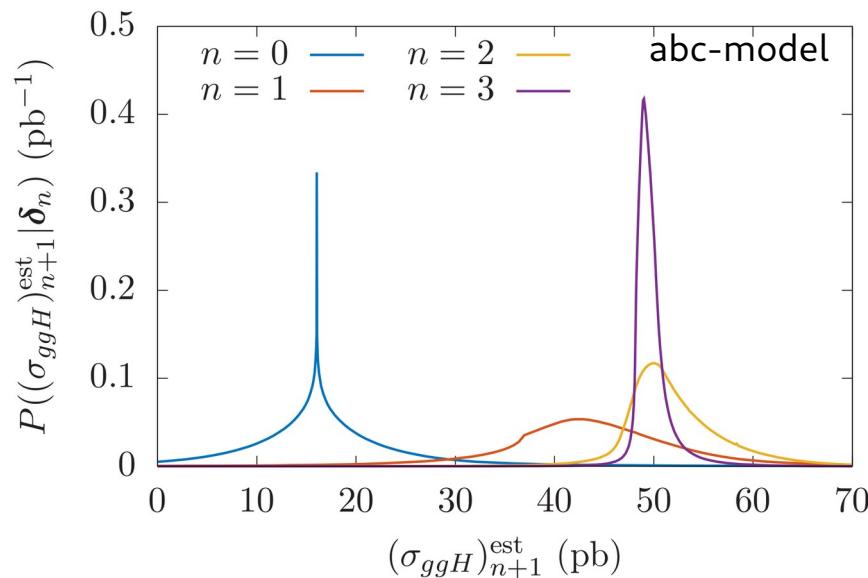
$$\mathcal{P}_{\text{sm}}(\Sigma|\Sigma_n) \approx \frac{\int d\mu P(\Sigma - \Sigma_n(\mu)|\Sigma_n(\mu)) P(\Sigma_n(\mu)) P_0(\mu)}{\int d\mu' P(\Sigma_n(\mu')) P_0(\mu')} . \quad \mu_{\text{FAC}}$$

Scale average (the results are averaged with weight function)

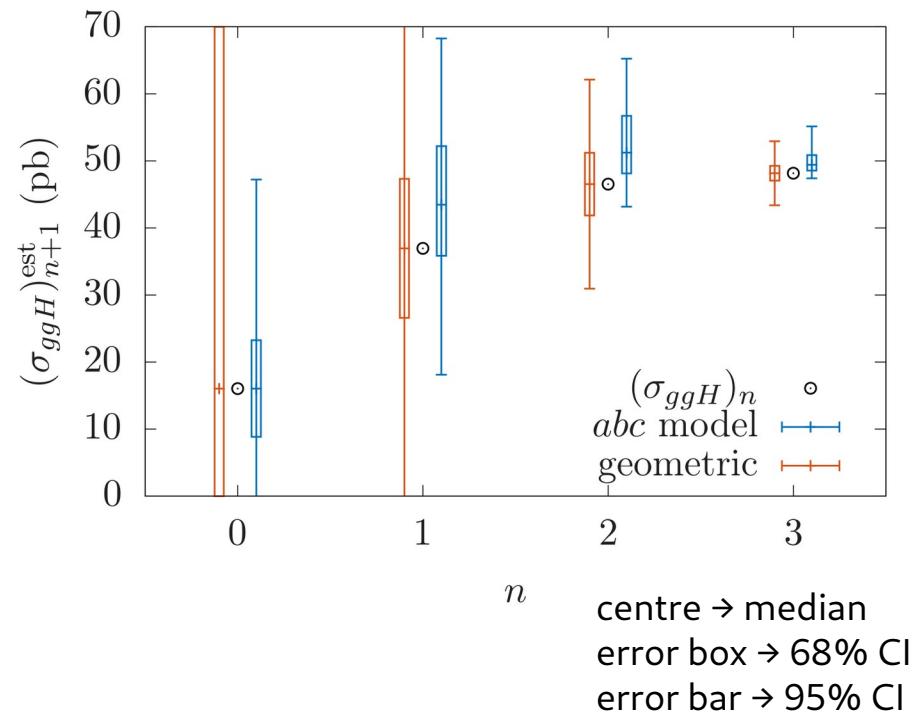
$$\mathcal{P}_{\text{sa}}(\Sigma|\Sigma_n) \approx \int d\mu w(\mu) P(\Sigma - \Sigma_n(\mu)|\Sigma_n(\mu)) \quad \mu_{\text{PMS}}$$

# Bayesian approach III

Example: Higgs production in gluon - fusion



Comparison of different unc. estimates:



# Introducing theory nuisance parameters (TNPs)

---

Generic perturbative expansion:

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

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

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

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

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

Introduce a parametrisation of unknown coefficients in terms of

"Theory nuisance parameters"  $\theta$

- The parametrization such that there is a true value:  $f_i(\hat{\theta}) = \hat{f}_i$
- Distributions of  $\theta$  "known" (for example from already existing computations)  
→ Expert knowledge to construct such a parametrisation

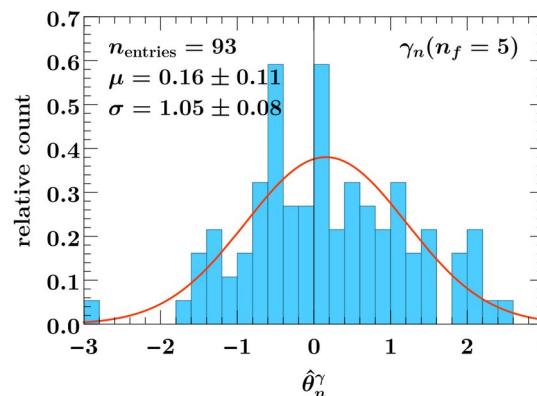
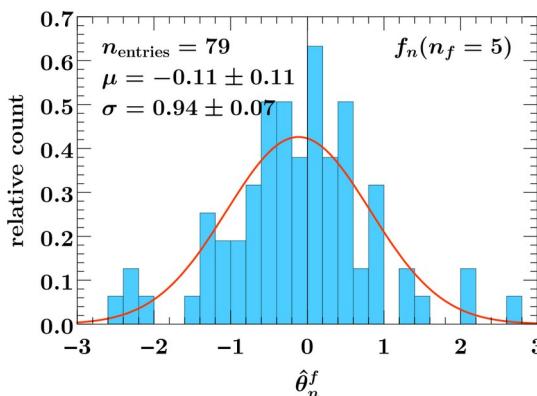
# Example: TNPs in resummed cross sections

Transverse momentum resummation:

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

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

Perturbative expansion of resummation ingredients:



$$X(\alpha_s) = X_0 + \alpha_s X_1 + \alpha_s^2 X_2 + \dots$$

$$\Gamma(\alpha_s) = \alpha_s [\Gamma_0 + \alpha_s \Gamma_1 + \alpha_s^2 \Gamma_2 + \dots]$$

$$\gamma(\alpha_s) = \alpha_s [\gamma_0 + \alpha_s \gamma_1 + \alpha_s^2 \gamma_2 + \dots]$$

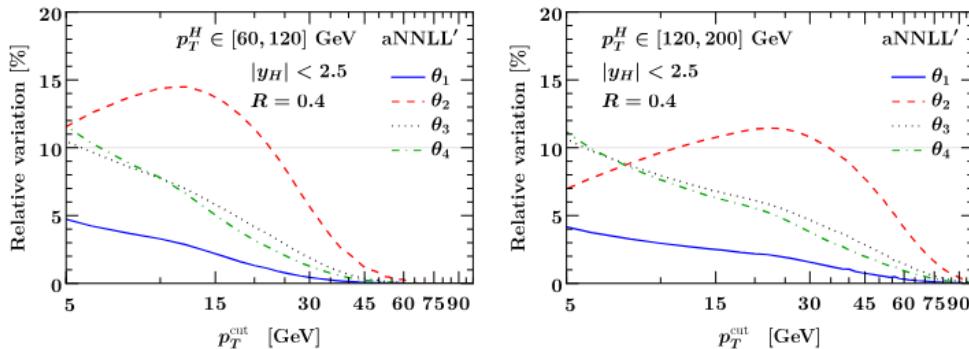
Normalisation:

$$f_n(\theta_n^f) = N_n^f \theta_n^f \quad \text{with} \quad N_n^f = 4^n C_r C_A^{n-1} (n-1)!$$

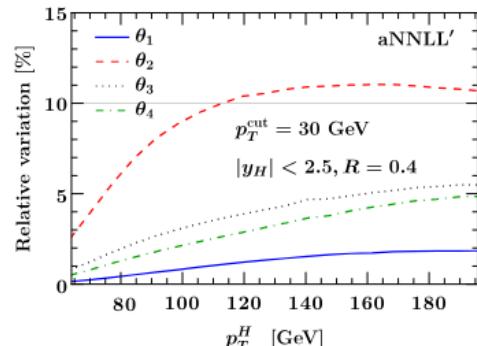
$$\gamma_n(\theta_n^\gamma) = N_n^\gamma \theta_n^\gamma \quad \text{with} \quad N_n^\gamma = 4^{n+1} C_r C_A^n$$

# Higgs pT spectrum

[Cal,Lim, Scott,Tackmann Waalewijn 2408.13301]



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



Example: incomplete knowledge of NNLL resummation

→ some two-loop ingredients unknown

→ parametrise by TNPs

→ Make predictions and vary TNPs:

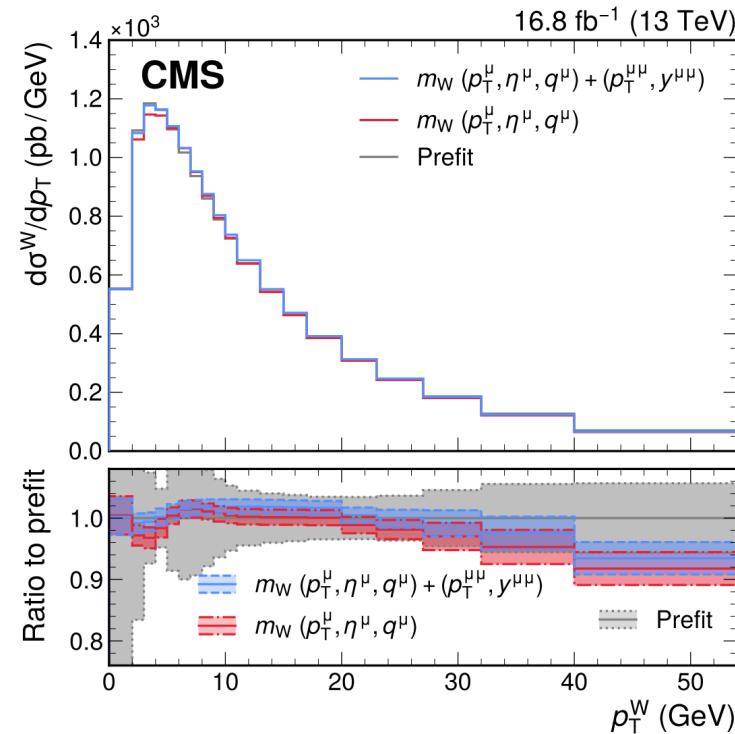
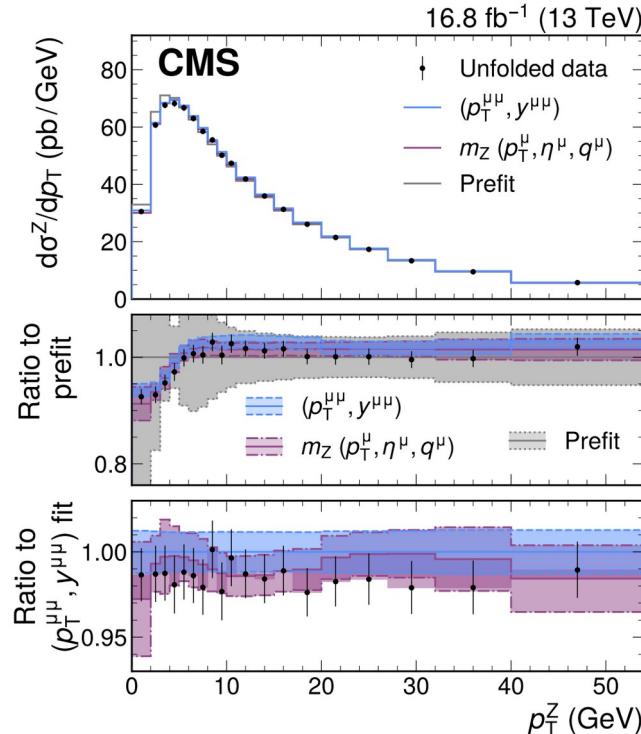
- correlated uncertainty for different bins!
- See impact of different missing ingredients

# Constraint of TNPs from data → W-mass extraction

Resummation ingredients the same for W or Z production

→ constraint from precisely measured  $p_T(Z)$  → use for  $p_T(W)$

→ 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,...)

# TNP approach for fixed-order computations

[Lim, Poncelet, 2412.14910]

$$d\sigma = \alpha_s^n N_c^m d\bar{\sigma}^{(0)} + \alpha_s^{n+1} N_c^{m+1} d\bar{\sigma}^{(1)} + \alpha_s^{n+2} N_c^{m+2} d\bar{\sigma}^{(2)} + \dots$$

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

Observation, i.e. "expert knowledge":  $\frac{d\bar{\sigma}^{(i)}}{d\bar{\sigma}^{(0)}} \sim \mathcal{O}(1)$

Use some knowledge about lower orders but introduce parametric dependence:

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

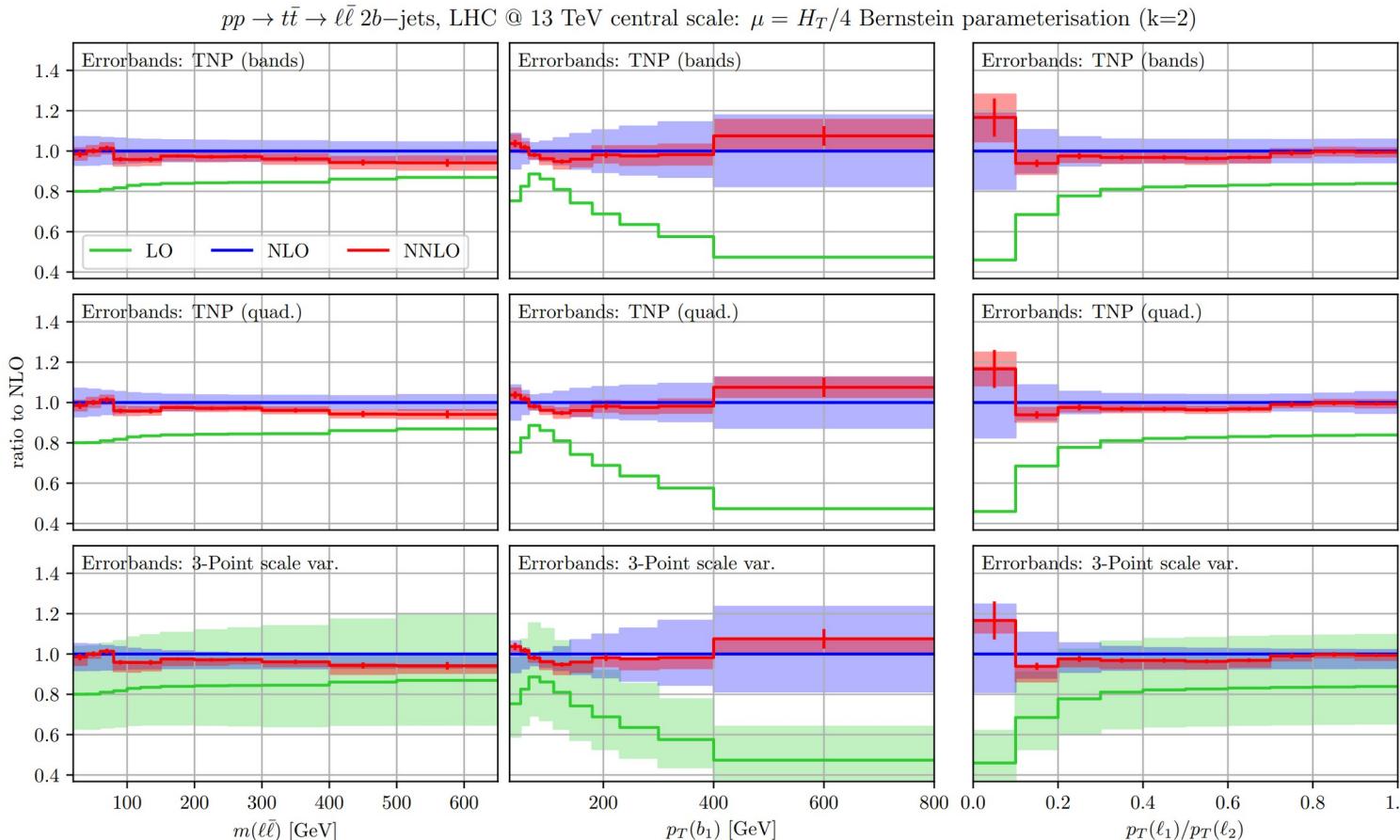
$x \rightarrow$  mapped kinematic variable

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

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

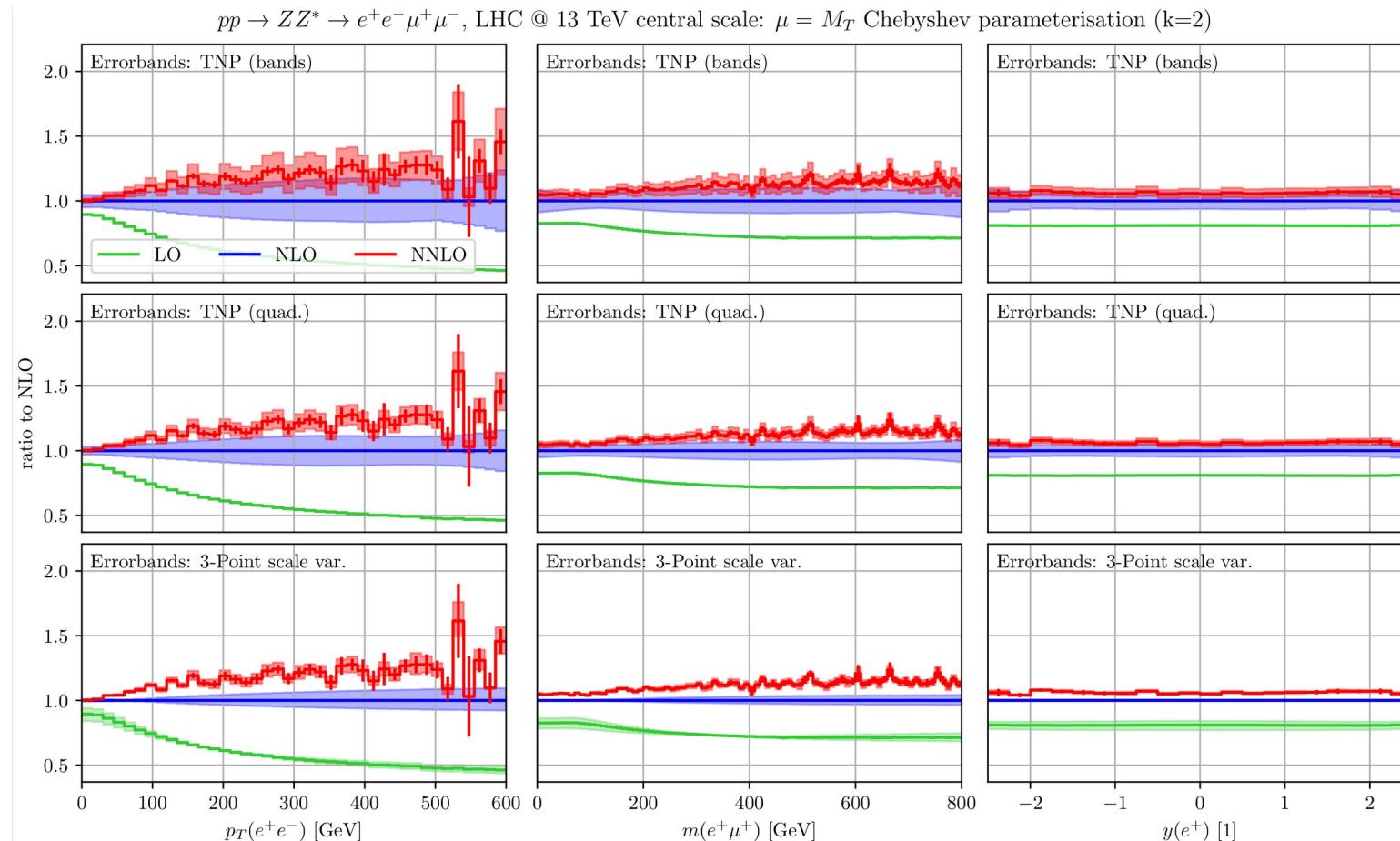
Chebyshev:  $f_k^C(\vec{\theta}, x) = \frac{1}{2} \sum_{i=0}^k \theta_i T_i(x)$   
 $x \in [-1, 1]$

# Uncertainties from TNPs - ttbar+decays



Band: sample  
 $\theta \in [-1, 1]$   
Quad: add individual  
 $\theta = \pm 1$   
in quadrature

# Uncertainties from TNPs - ZZ

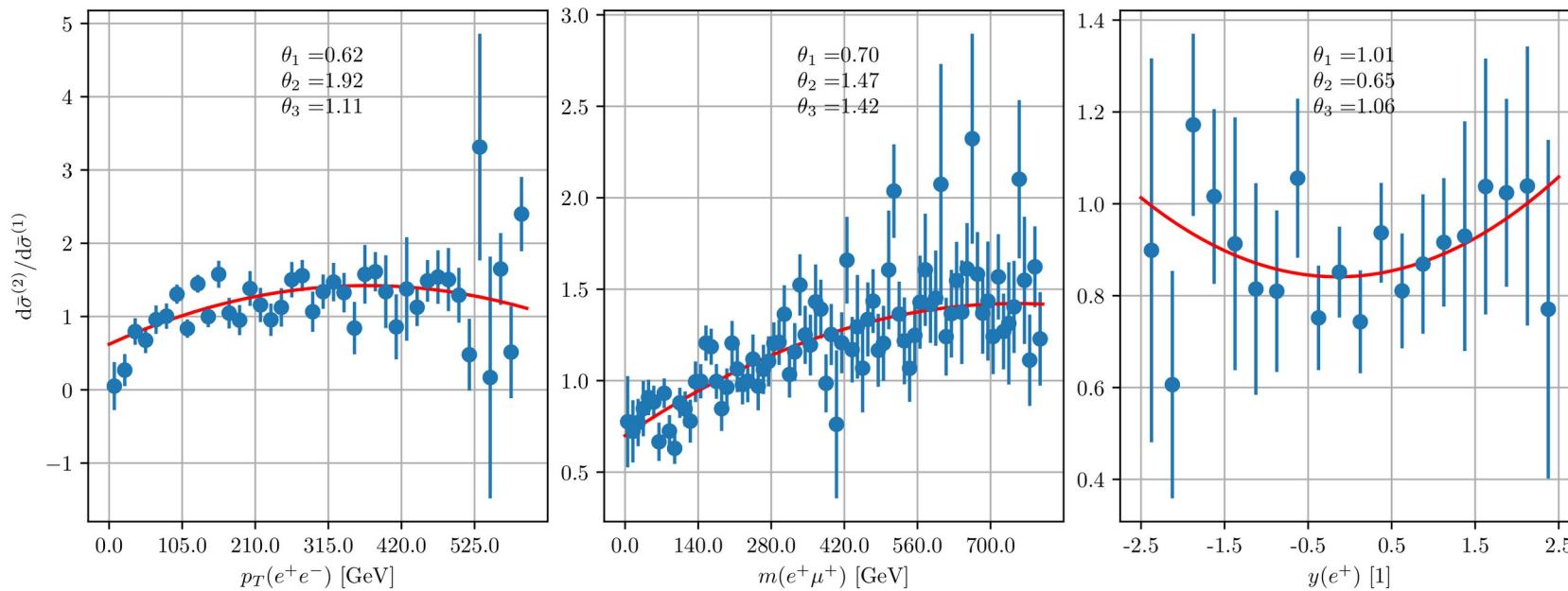


# Example of TNP fit: $pp \rightarrow ZZ$

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

$$f_k^B(\vec{\theta}, x) = \sum_{i=0}^k \theta_i \binom{k}{i} x^{k-i} (1-x)^i$$

$pp \rightarrow ZZ^* \rightarrow e^+e^-\mu^+\mu^-$ , LHC @ 13 TeV central scale:  $\mu = M_T$



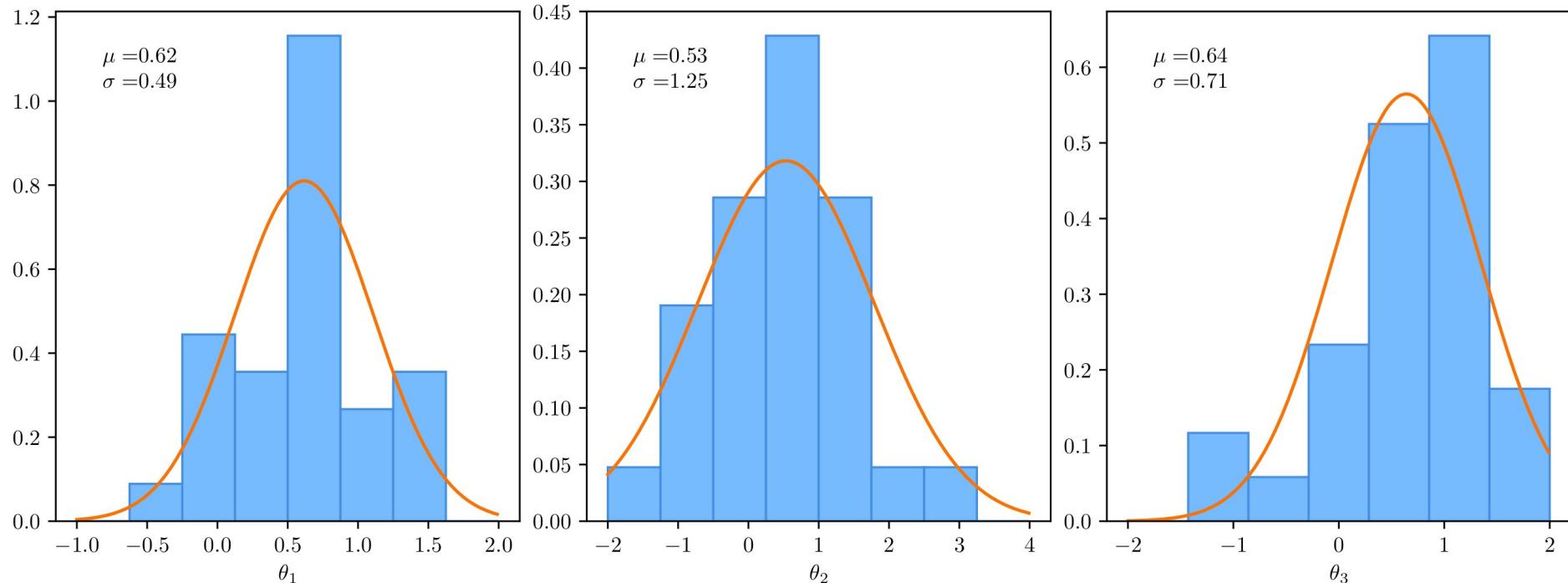
# Process meta study

---

Process	$\sqrt{s}/\text{TeV}$	Scale	PDF	Distributions
$pp \rightarrow H$ (full theory)	13	$m_H/2$	NNPDF3.1	$y_H$
$pp \rightarrow ZZ^* \rightarrow e^+e^-\mu^+\mu^-$	13	$M_T$	NNPDF3.1	$M_{e^+\mu^+}, p_T^{e^+e^-}, y_{e^+}$
$pp \rightarrow WW^* \rightarrow e\nu_e\mu\nu_\mu$	13	$m_W$	NNPDF3.1	$M_{WW}, p_T^{\mu^-}, y_{W^-}$
$pp \rightarrow (W \rightarrow \ell\nu) + c$	13	$E_T + p_T^c$	NNPDF3.1	$p_T^\ell,  y_\ell ,$
$pp \rightarrow t\bar{t}$	13	$H_T/4$	NNPDF3.1	$M_{t\bar{t}}, p_T^t, y_t$
$pp \rightarrow t\bar{t} \rightarrow b\bar{b}\ell\bar{\ell}$	13	$H_T/4$	NNPDF3.1	$M_{\ell\bar{\ell}}, p_T^{b_1}, p_{T,\ell_1}/p_{T,\ell_2}$
$pp \rightarrow \gamma\gamma$	8	$M_{\gamma\gamma}$	NNPDF3.1	$M_{\gamma\gamma}, p_T^{\gamma_1}, y_{\gamma\gamma}$
$pp \rightarrow \gamma\gamma j$	13	$H_T/2$	NNPDF3.1	$M_{\gamma\gamma}, p_T^{\gamma\gamma}, \cos\phi_{\text{CS}},  y_{\gamma_1} , \Delta\phi_{\gamma\gamma}$
$pp \rightarrow jjj$	13	$\hat{H}_T$	MMHT2014	TEEC with $H_{T,2} \in [1000, 1500), [1500, 2000), [3500, \infty)$ GeV
$pp \rightarrow \gamma jj$	13	$H_T$	NNPDF3.1	$M_{\gamma jj}, p_T^j,  y_{\gamma-\text{jet}} , E_{T,\gamma}$

# Fits - Bernstein parametrisation

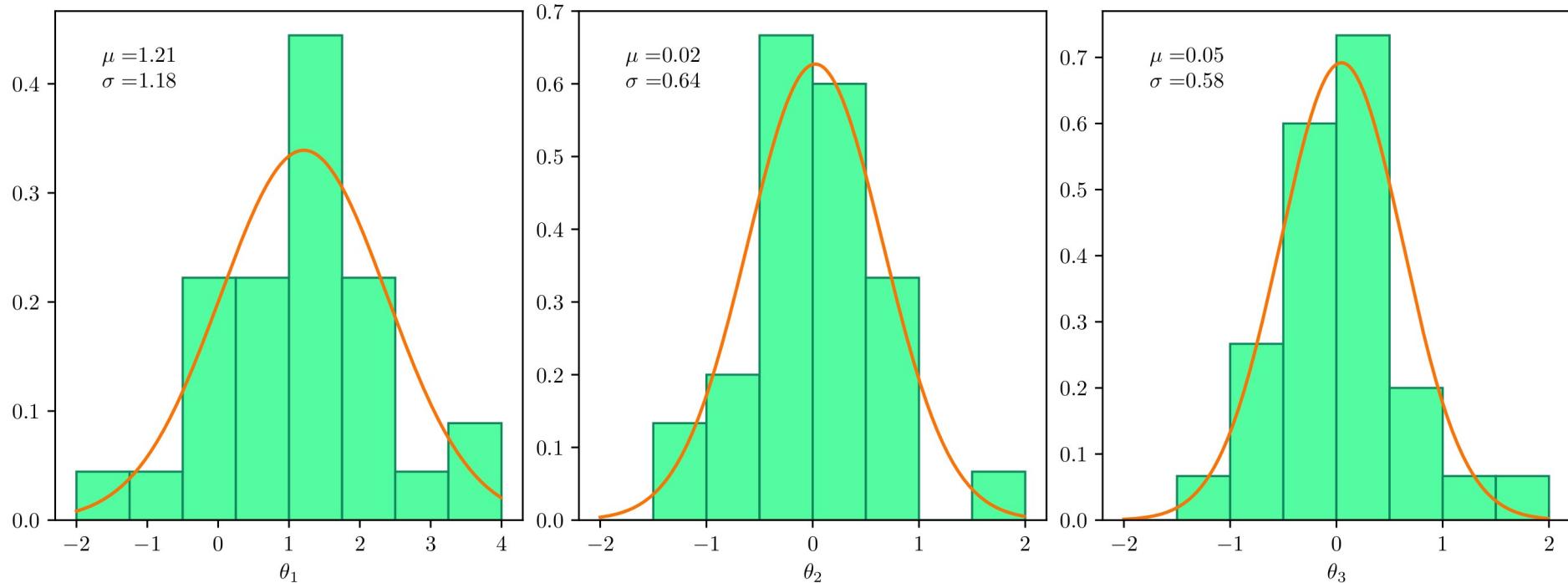
TNPs in Bernstein parameterisation



$$f_k^B(\vec{\theta}, x) = \sum_{i=0}^k \theta_i \binom{k}{i} x^{k-i} (1-x)^i$$

# Fits - Chebyshev parametrisation

TNPs in Chebyshev parameterisation



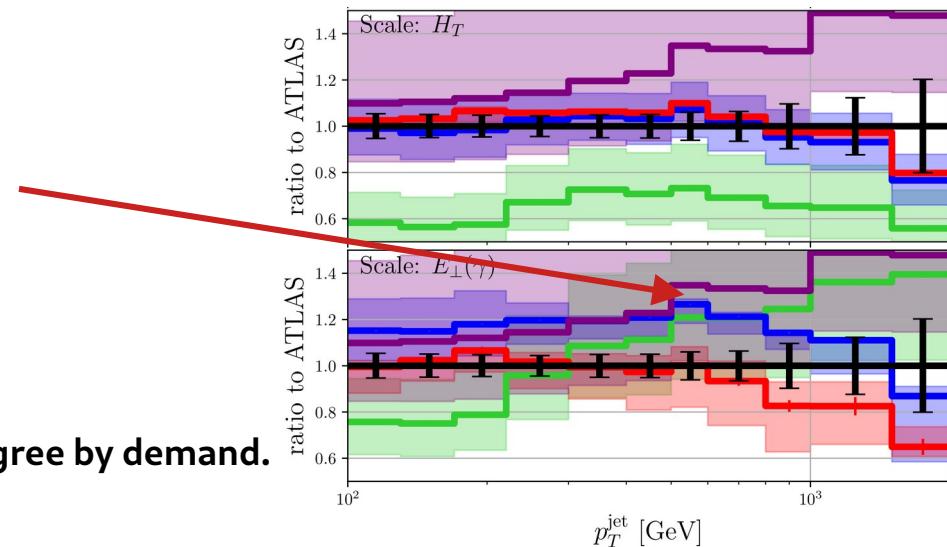
$$f_k^C(\vec{\theta}, x) = \frac{1}{2} \sum_{i=0}^k \theta_i T_i(x)$$

# Caveats and open questions

Some arising questions regarding fixed-order model:

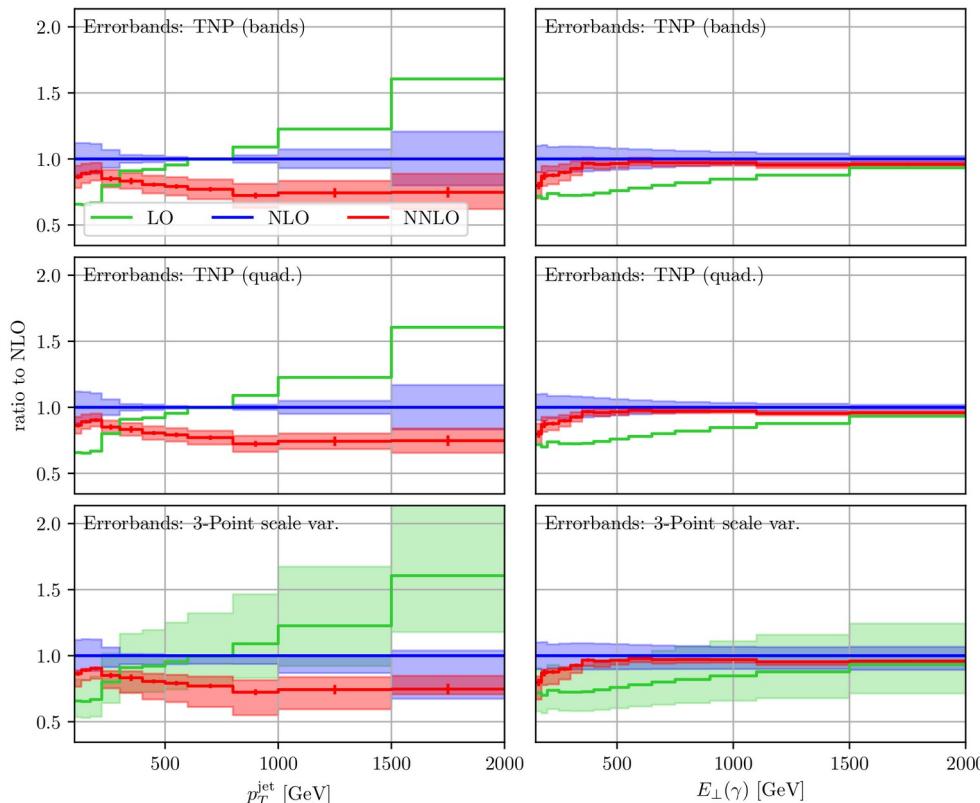
- How does the uncertainty estimate depend on the central scale choice?  
→ **bad scale choices lead to large uncertainties by construction due to large corrections.**
- What about NLO uncertainty if  $d\bar{\sigma}^{(1)} = 0$  for given scale?  
→ **amend parametrisation by j = 0 term.**
- How sensitive are we to the parametrisation?  
How many terms?  
→ **two quite general parametrisations tested, increase degree by demand.**

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

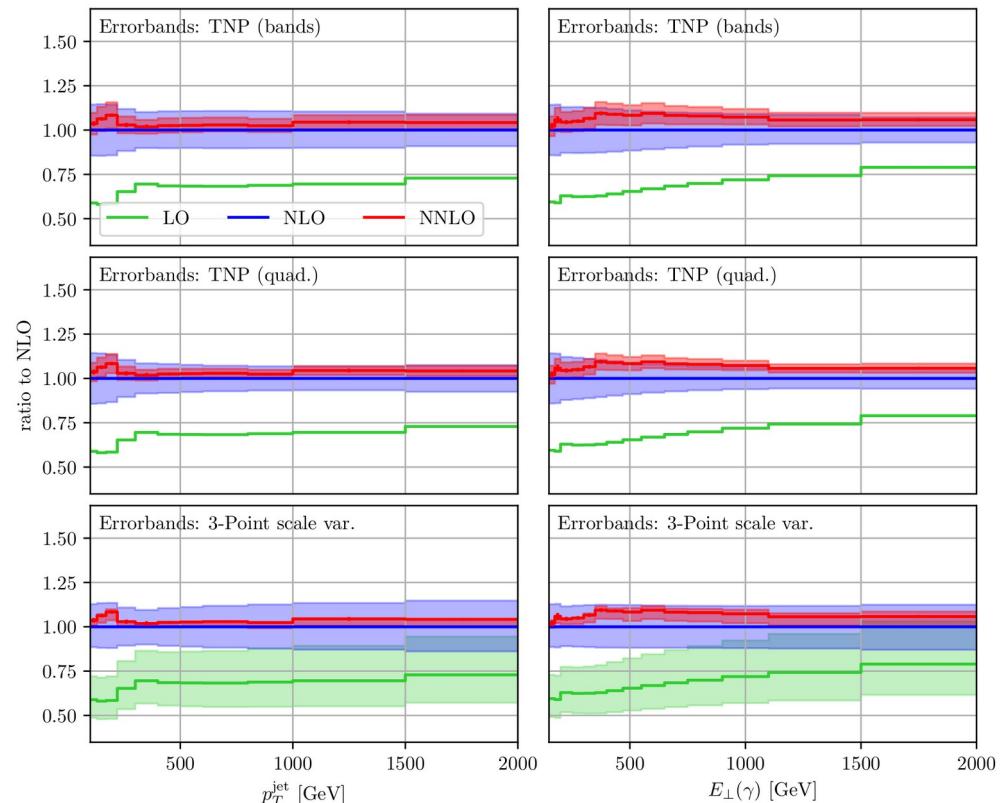


# Challenging case

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

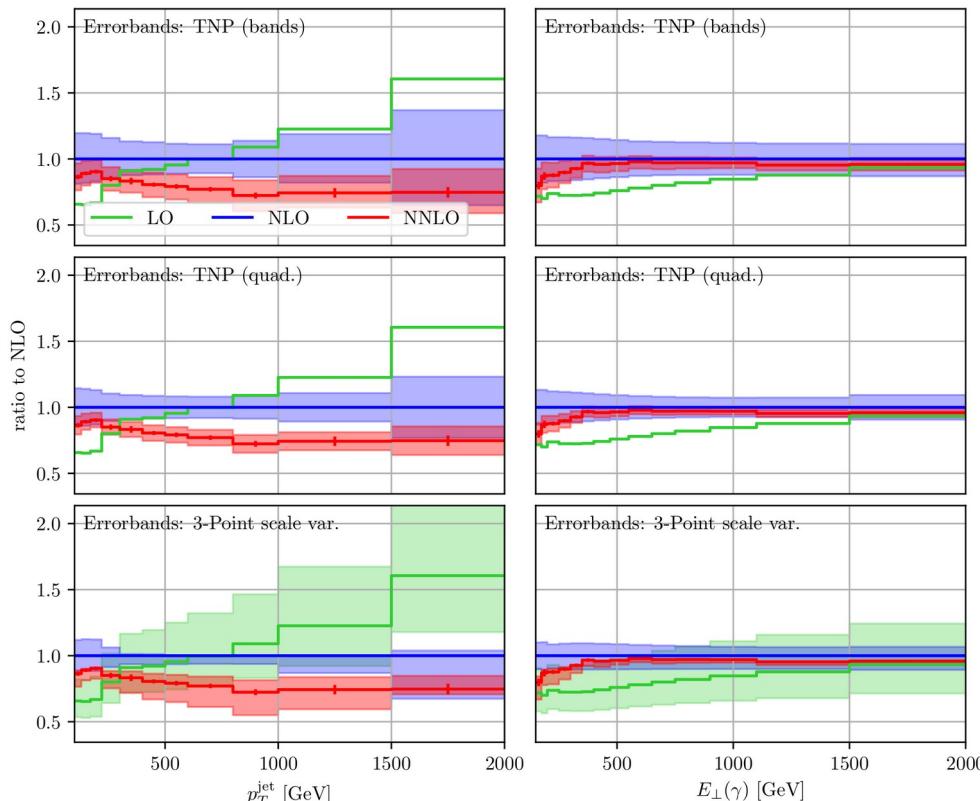


“Good” choice  $\mu = H_T$  no  $j=0$  term

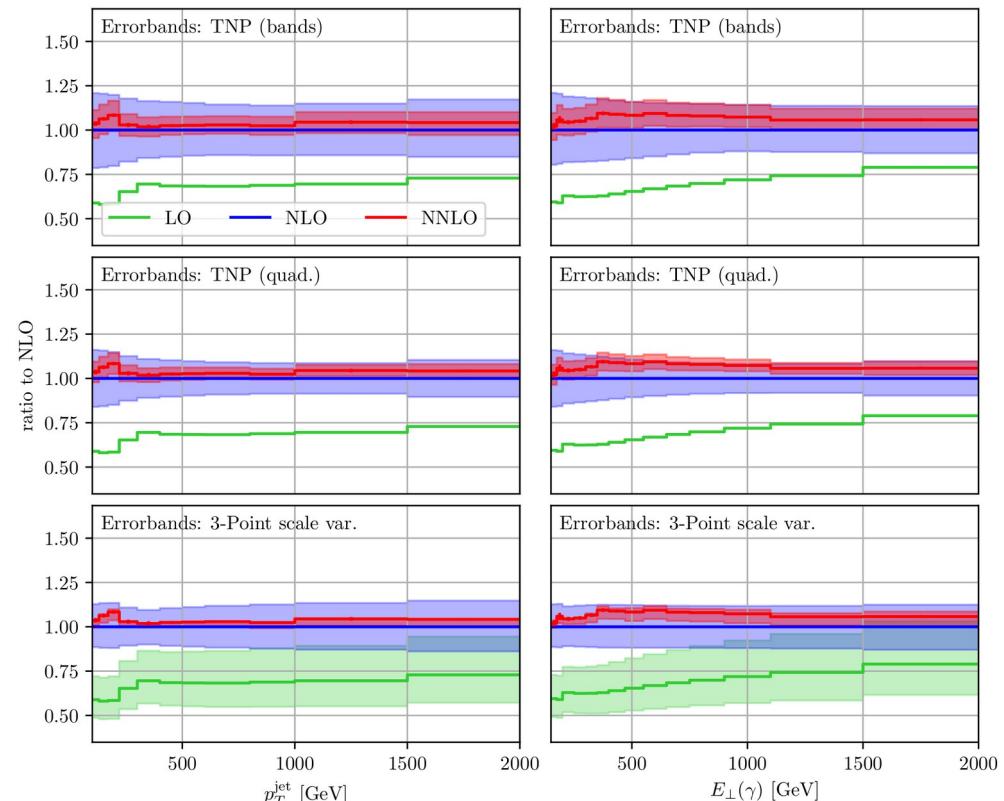


# Challenging case → extended parametrisations

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

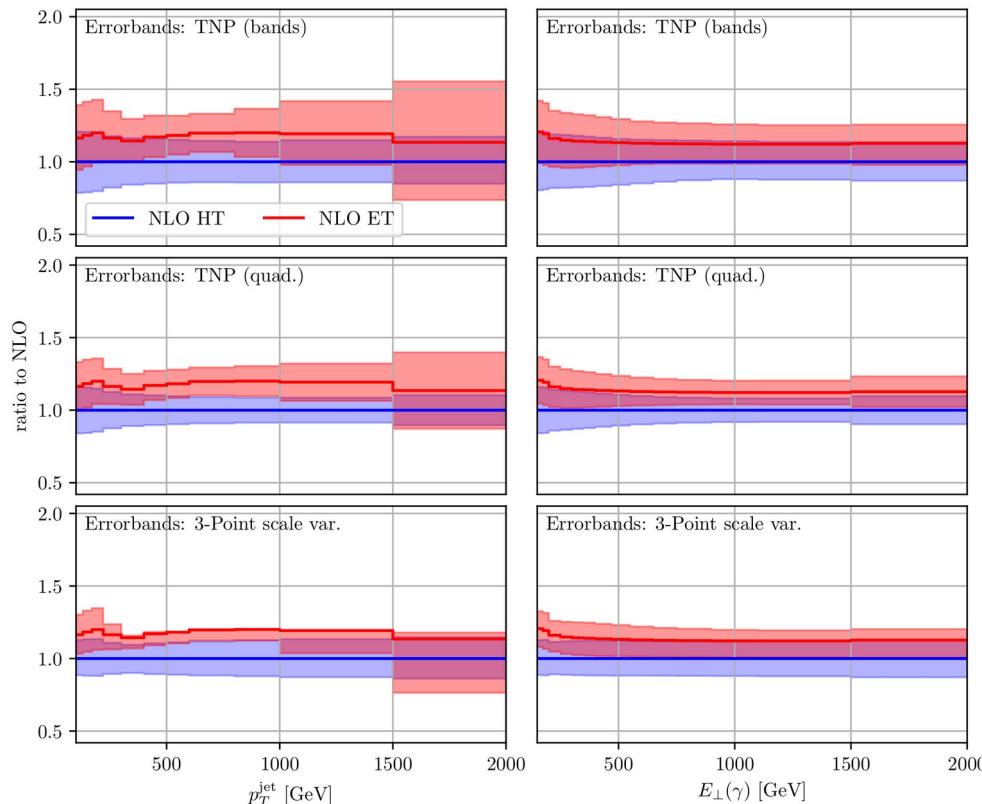


“Good” choice  $\mu = H_T$  with j=0 term

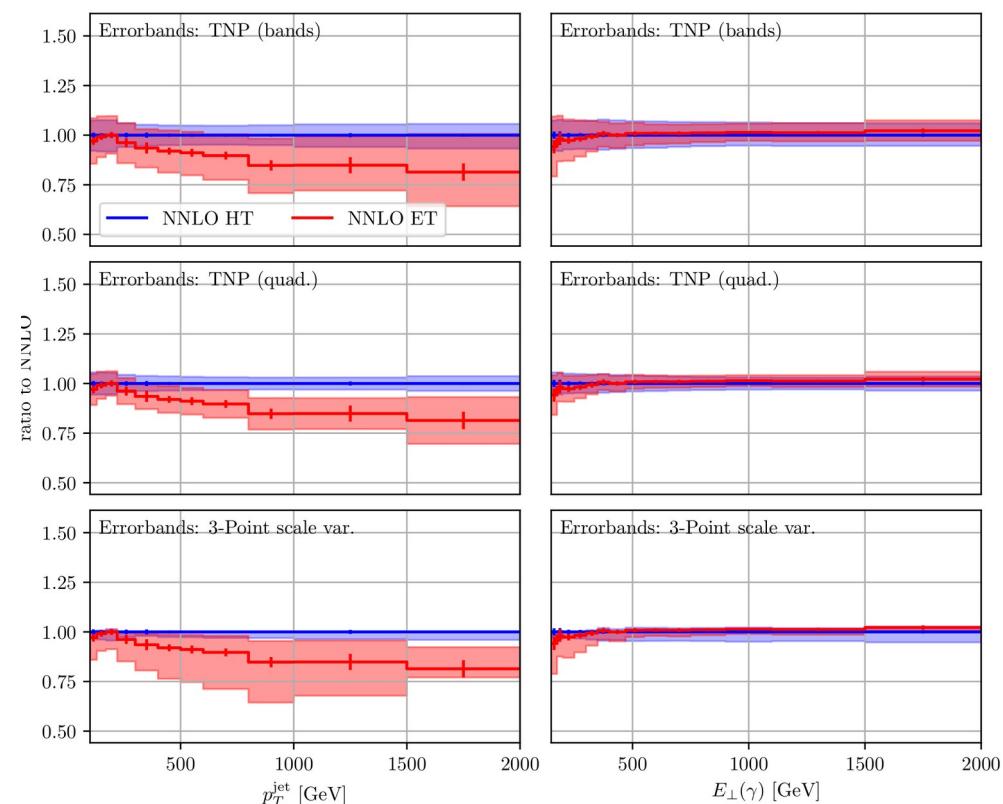


# Challenging case → comparisons

NLO QCD



NNLO QCD



# More caveats and open questions

---

Some arising questions regarding fixed-order model:

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

- Each parametrisation is for one observable at a time
  - How to deal with higher dimensional distributions?
  - Consistency upon integration?  
→ work in progress
- What about EW corrections?  
→ here the approach should work well for Sudakov logs!  
→ Radiation from resonances more difficult.
- How to correlate different processes?  
→ that's tricky...

$$d\bar{\sigma}^{(n)}(\theta) = d\Phi \langle M^0 | \mathcal{P}(\theta) | M^0 \rangle$$

$\mathcal{P}(\theta)$  → process-independent “operator”

# Discussion/Summary/Outlook

---

- Theory precision is more and more relevant → needs accurate 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 → does not depend on any “known” orders
  - Statistical interpretation
  - Needs “expert knowledge”
  - Resummed computations are well-suited scenario
  - Fixed-order more difficult: not much knowledge about higher-order terms  
Proposed TNP parametrisation shows promising results  
→ applications to fits