

Theory uncertainties from theory nuisance parameters in fixed-order QCD (online)

Rene Poncelet

based on [[Lim, Poncelet, 2412.14910](#)]

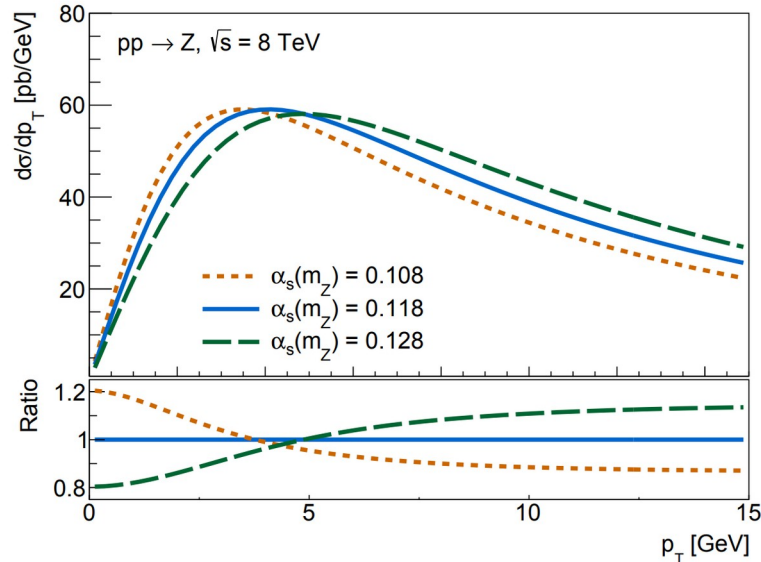


THE HENRYK NIEWODNICZAŃSKI
INSTITUTE OF NUCLEAR PHYSICS
POLISH ACADEMY OF SCIENCES

Precision example: strong coupling from pT(Z)

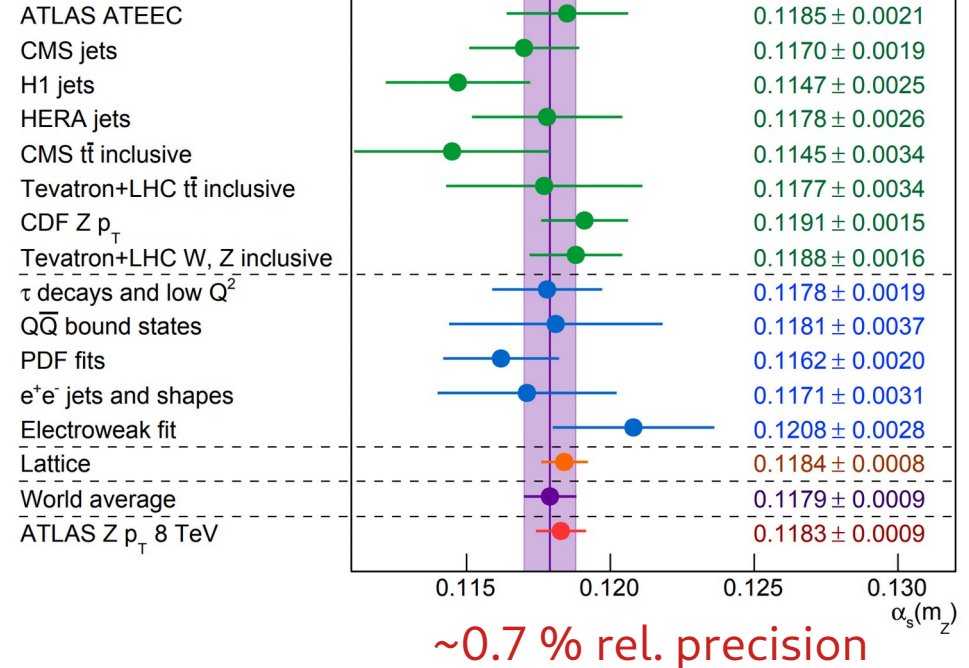
[ATLAS 2309.12986]

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



→ at low pT resummation regime!

→ theory uncertainty?

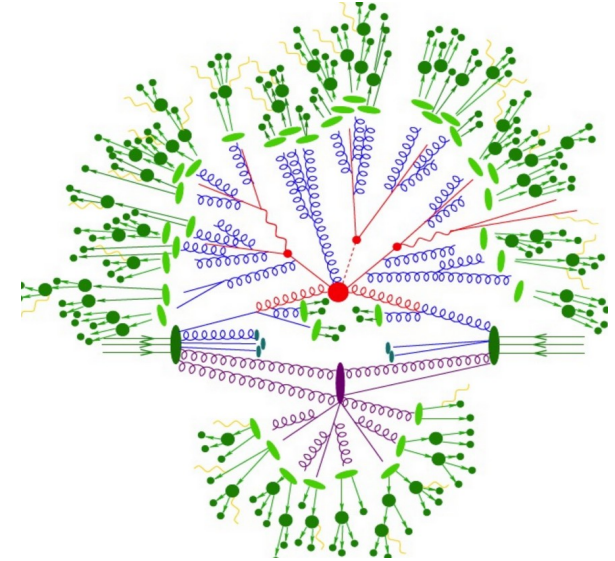


Uncertainties in precision phenomenology

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

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

The missing terms are the source of uncertainty.

(assume convergence $\rightarrow$ the first missing is the dominant one)

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

Challenge: how to estimate $f_1, f_2, f_3, \dots$ without computing them?

Theory uncertainties from scale variations

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

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

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

Scale variation prescription (ad-hoc and heuristic choice)

- choose 'sensible' μ_0

- $\rightarrow$ principle of fastest apparent convergence:

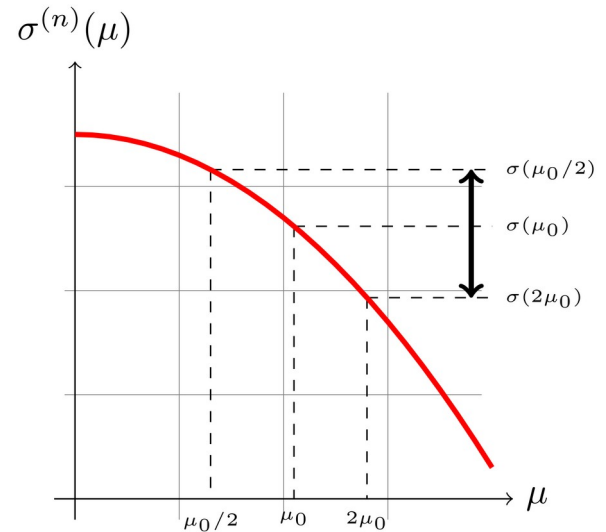
- $\rightarrow$ principle of minimal sensitivity

- $\rightarrow \dots$

- vary with a factor (typically 2)

- take envelope as uncertainty

$$\sigma^{(n)}(\mu_{\text{FAC}}) = 0$$
$$\mu \frac{d\sigma(\mu)}{d\mu} \Big|_{\mu=\mu_{\text{PMC}}} = 0$$



Scale variation approach

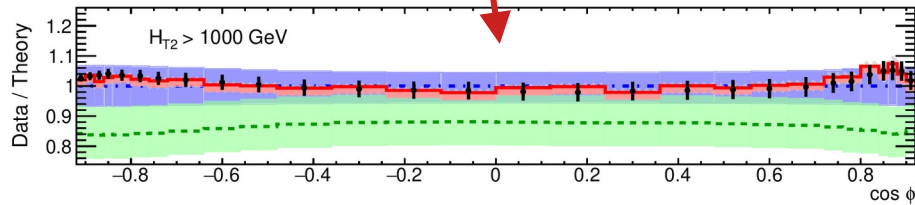
Estimates from scale variations for the unknown f_i :

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

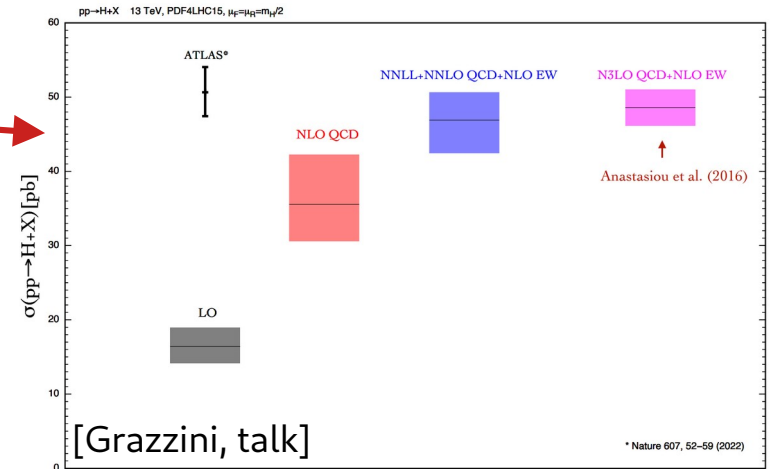
$$f_3 \equiv \Delta f^{\text{NNLO}} = f^{\text{NNLO}} - \tilde{f}^{\text{NNLO}}(\tilde{\alpha}) = \alpha^3 (2b_0(\hat{f}_2 - b_0 \hat{f}_1) + b_1 \hat{f}_1) + \mathcal{O}(\alpha^4)$$

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

Sometimes it does work ... sometimes not



[ATLAS 2301.09351]



[Grazzini, talk]

* Nature 607, 52–59 (2022)

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]

[Cridge,Marinelli,Tackmann,2506.13874]

Introducing theory nuisance parameters (TNPs)

[Tackman 2411.18606]

Generic perturbative expansion:

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

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

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

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

Introduce a parametrisation of unknown coefficients in terms of

"Theory nuisance parameters" 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:

$$\frac{d\sigma}{dp_T} = \underbrace{[H \times B_a \times B_b \times S]}_{X(\alpha_s, L)}(\alpha_s, \ln\left(\frac{p_T}{Q}\right)) + \mathcal{O}\left(\frac{p_T}{Q}\right)$$
$$X(\alpha_s, L) = X(\alpha_s) \exp \int_0^L dL' \{ \Gamma(\alpha_s(L')) L' + \gamma_X(\alpha_s(L')) \}$$

Perturbative expansion of resummation ingredients:

$$X(\alpha_s) = X_0 + \alpha_s X_1 + \alpha_s^2 X_2 + \cdots + \alpha_s^n X_n(\theta)$$
$$\Gamma(\alpha_s) = \alpha_s [\Gamma_0 + \alpha_s \Gamma_1 + \alpha_s^2 \Gamma_2 + \cdots + \alpha_s^n \Gamma_n(\theta)]$$
$$\gamma(\alpha_s) = \alpha_s [\gamma_0 + \alpha_s \gamma_1 + \alpha_s^2 \gamma_2 + \cdots + \alpha_s^n \gamma_n(\theta)]$$

Task: find suitable parametrization and variation range

These are numbers for simple processes → only need normalisation

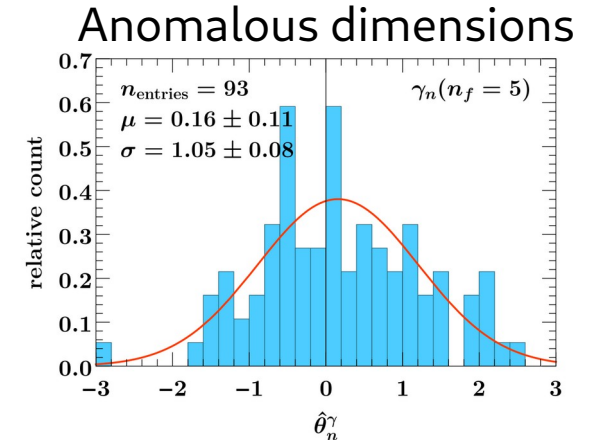
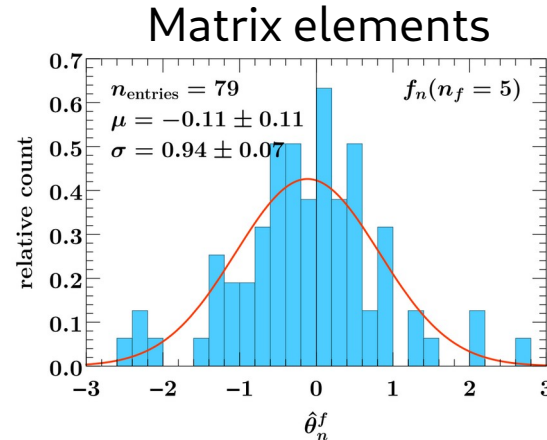
TNP parametrisations for resummation

[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
	$4^{n+1}C_FC_A^n$	-1.28	-0.54	-0.21	-0.47	-0.12
γ_m	1	-8.00	-112	-950	-5650	-85648
	4^{n+1}	-2.00	-7.028	-14.8	-22.1	-83.6
	$4^{n+1}C_FC_A^n$	-1.50	-1.76	-1.24	-0.61	-0.77
$2\Gamma_{\text{cusp}}^q$	1	+10.7	+73.7	+478	+282	(+140000)
	4^{n+1}	+2.67	+4.61	+7.48	+1.10	(+137)
	$4^{n+1}C_FC_A^n$	+2.00	+1.15	+0.62	+0.03	(+1.27)

“Statistic of many computations”

•
•
•



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

$\rightarrow$ there is only $f_i(\hat{\theta}) \approx \hat{f}_i$

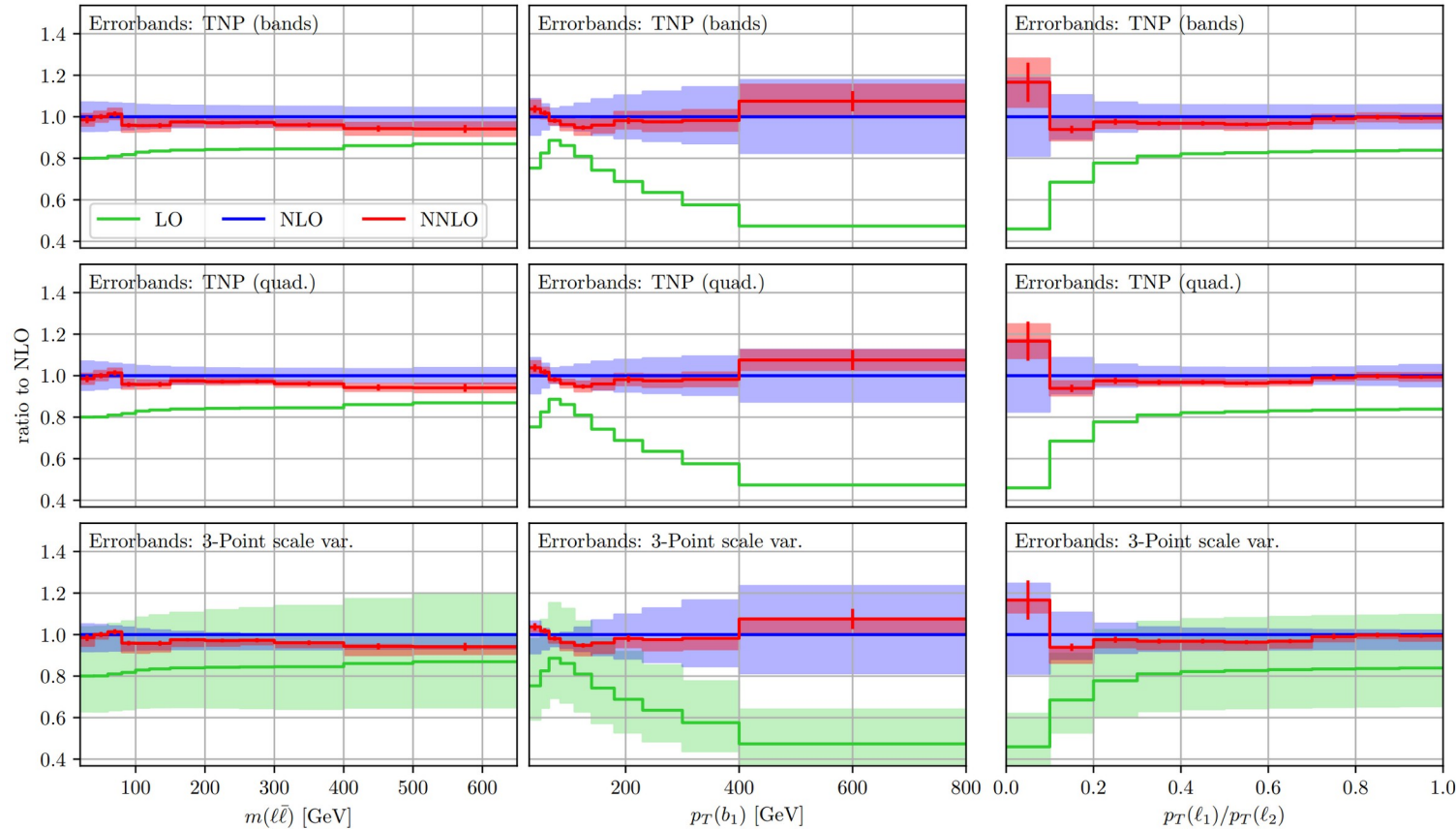
See also MSHT [2207.04739]

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

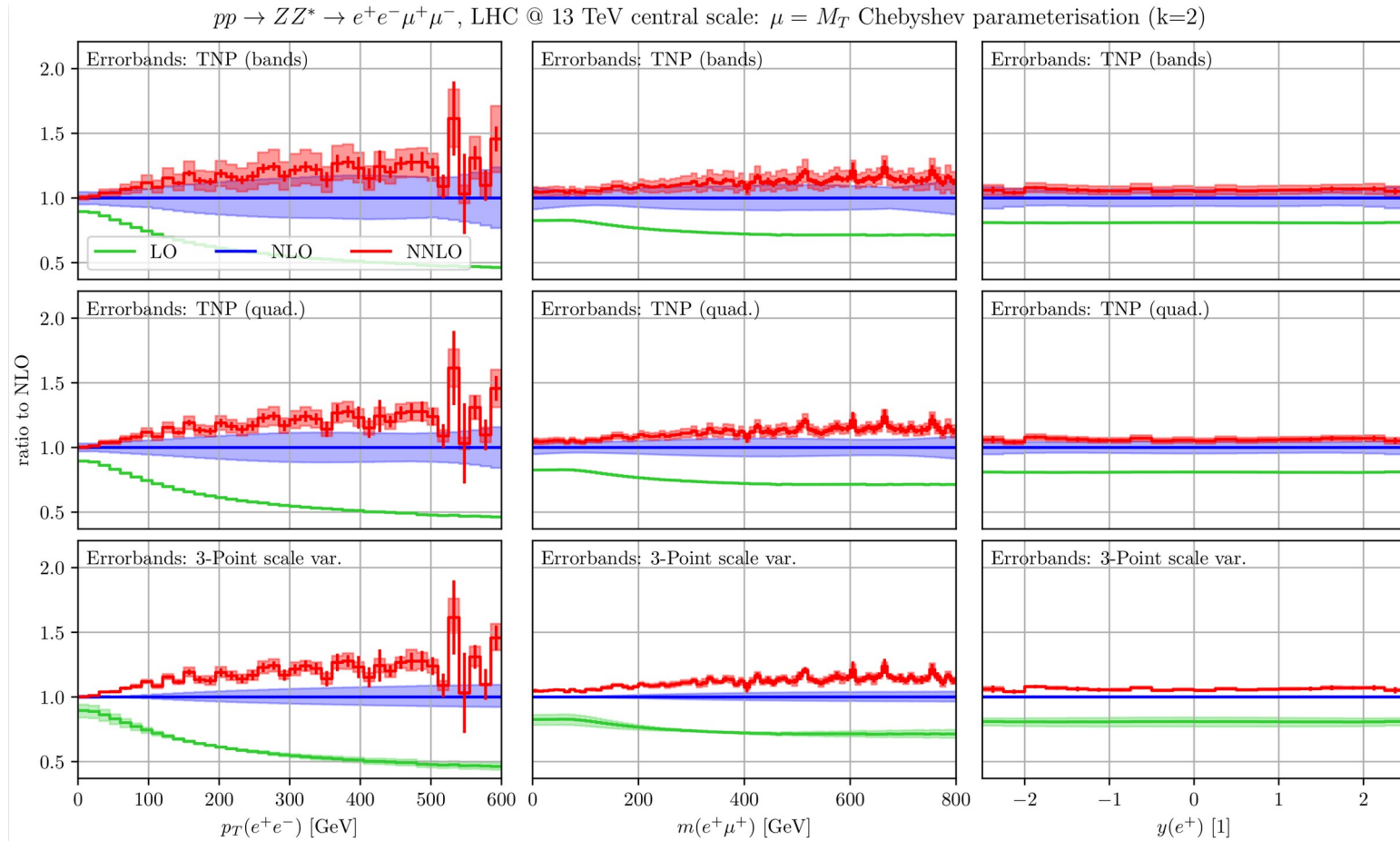
Uncertainties from TNP's - $t\bar{t}$ bar+decays

$pp \rightarrow t\bar{t} \rightarrow \ell\bar{\ell} 2b$ -jets, LHC @ 13 TeV central scale: $\mu = H_T/4$ Bernstein parameterisation (k=2)



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

Uncertainties from TNP's - ZZ

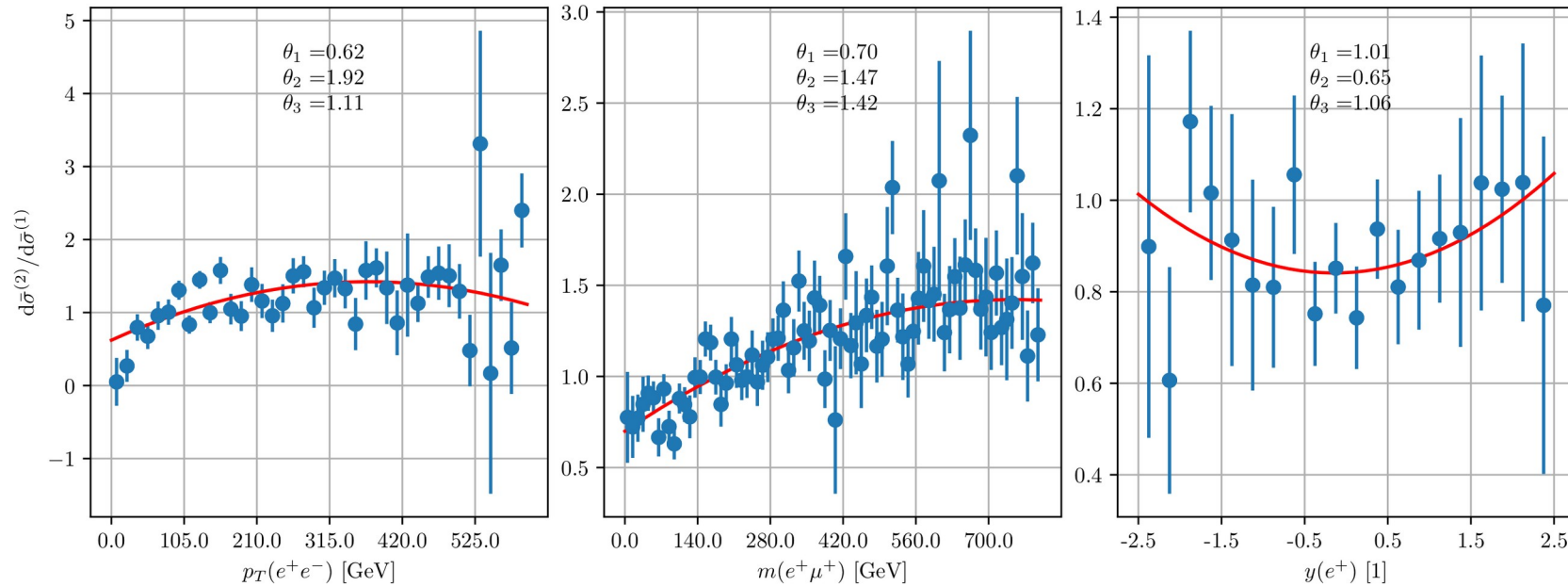


Example of TNP fit: $pp \rightarrow ZZ$

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

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

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

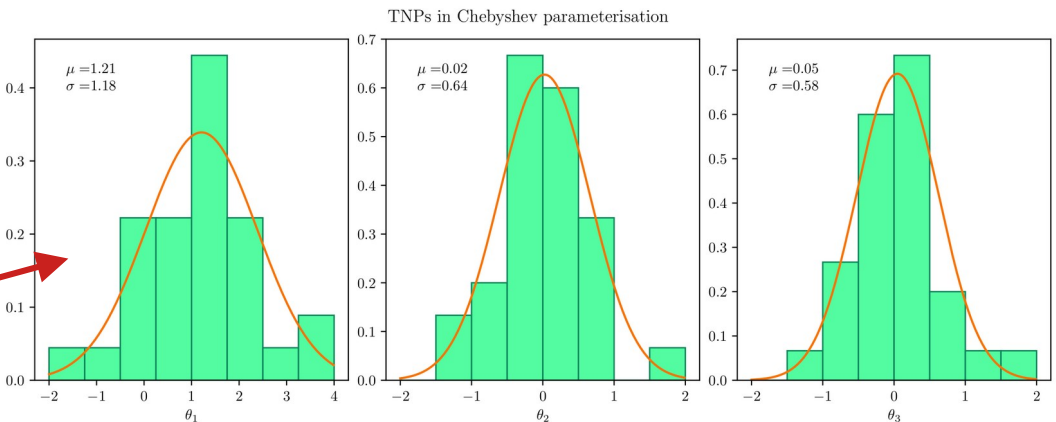
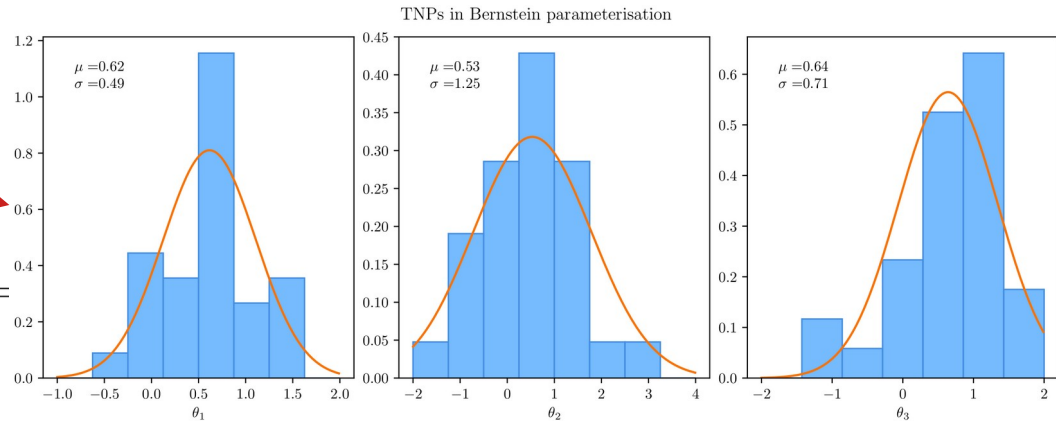


Fits - meta-study

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

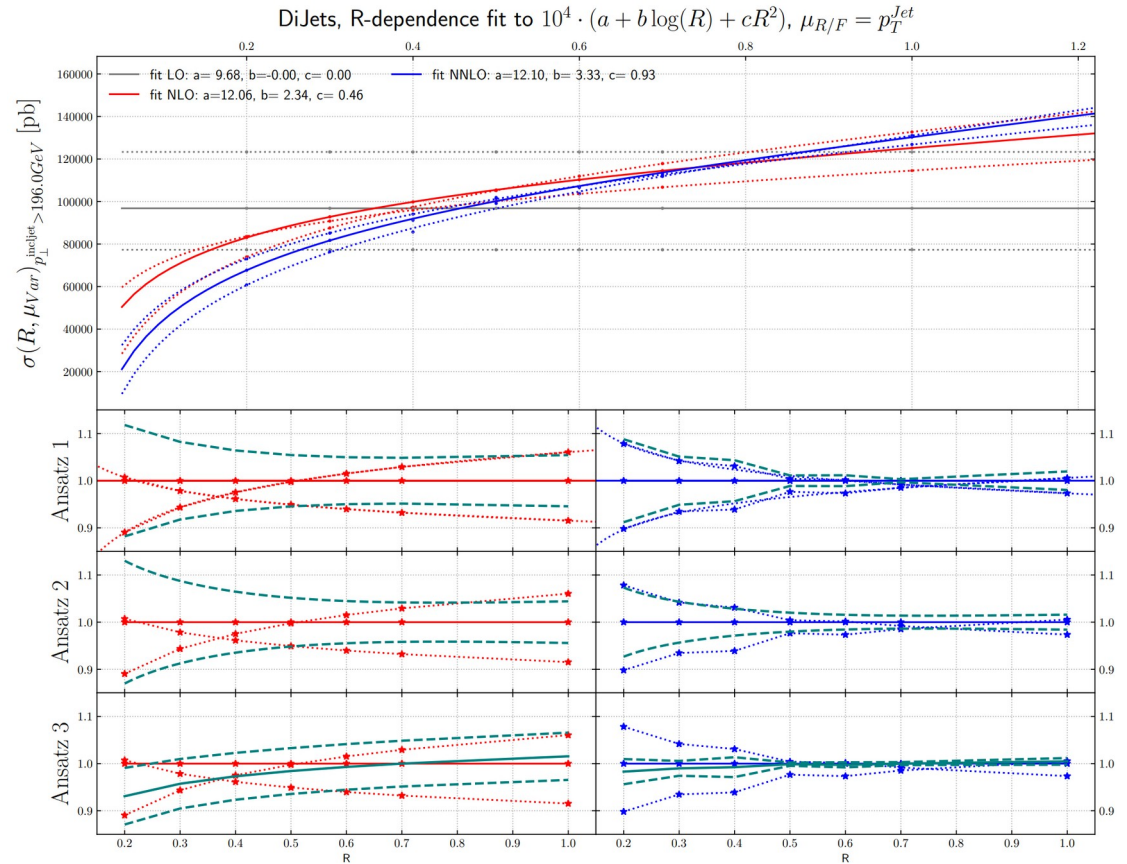
Process	Distributions
$pp \rightarrow H$ (full theory)	y_H
$pp \rightarrow ZZ^* \rightarrow e^+e^-\mu^+\mu^-$	$M_{e^+\mu^+}, p_T^{e^+e^-}, y_{e^+}$
$pp \rightarrow WW^* \rightarrow e\nu_e\mu\nu_\mu$	$M_{WW}, p_T^{\mu^-}, y_{W^-}$
$pp \rightarrow (W \rightarrow \ell\nu) + c$	$p_T^\ell, y_\ell $
$pp \rightarrow t\bar{t}$	$M_{t\bar{t}}, p_T^t, y_t$
$pp \rightarrow t\bar{t} \rightarrow b\bar{b}\ell\bar{\ell}$	$M_{\ell\bar{\ell}}, p_T^{b1}, p_{T,\ell1}/p_{T,\ell2}$
$pp \rightarrow \gamma\gamma$	$M_{\gamma\gamma}, p_T^{\gamma1}, y_{\gamma\gamma}$
$pp \rightarrow \gamma\gamma j$	$M_{\gamma\gamma}, p_T^{\gamma\gamma}, \cos\phi_{CS}, y_{\gamma1} , \Delta\phi_{\gamma\gamma}$
$pp \rightarrow jjj$	EEC with $H_{T,2} \in [1000, 1500), [1500, 2000), [3500, \infty)$
$pp \rightarrow \gamma jj$	$M_{\gamma jj}, p_T^j, y_{\gamma-jet} , E_{T,\gamma}$

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



Example: inclusive jet production

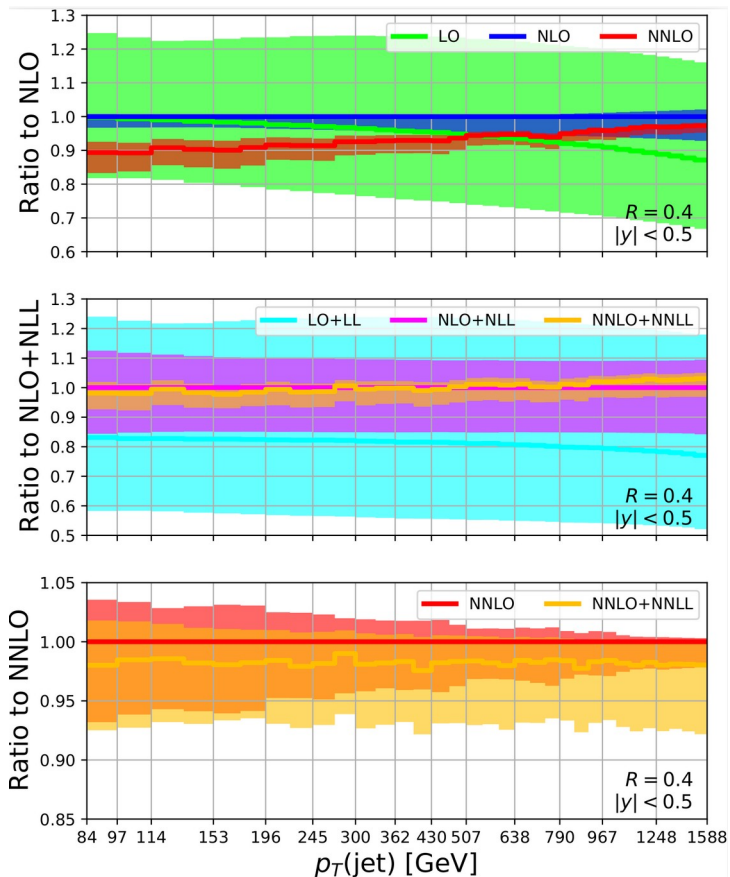
- Important process for PDF fits: sensitivity to gluon PDF at large-x
- NNLO QCD corrections imply very small theory uncertainty
- Significant jet radius dependence of uncertainties from scale variations



[1903.12563 Bellm et al]

Inclusive jet production: small-R resummation NNLO+NNLL

[Generet, Lee, Mout, Poncelet, Zhang'25]



FO scale variations

$R=0.4$

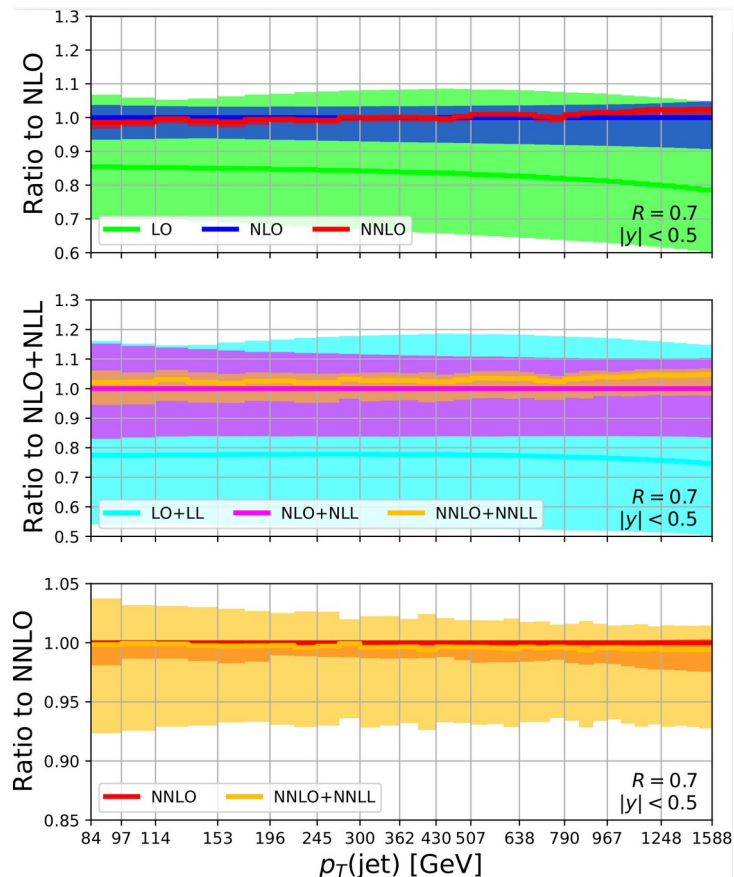
→ underestimation of NNLO correction

$R=0.7$

→ very small NNLO uncertainty

Resummation

→ stabilization of pert. series and uncertainties.

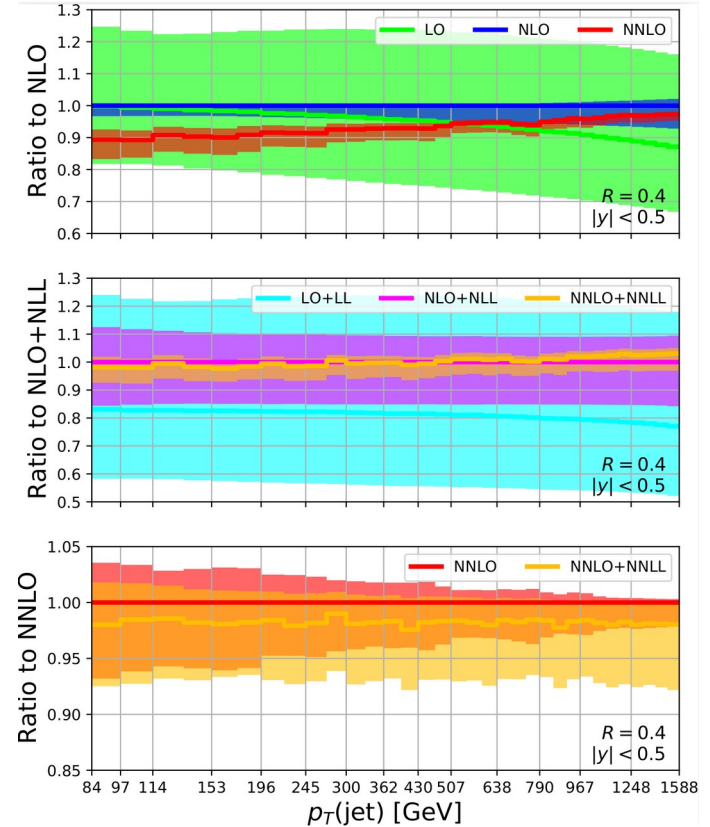
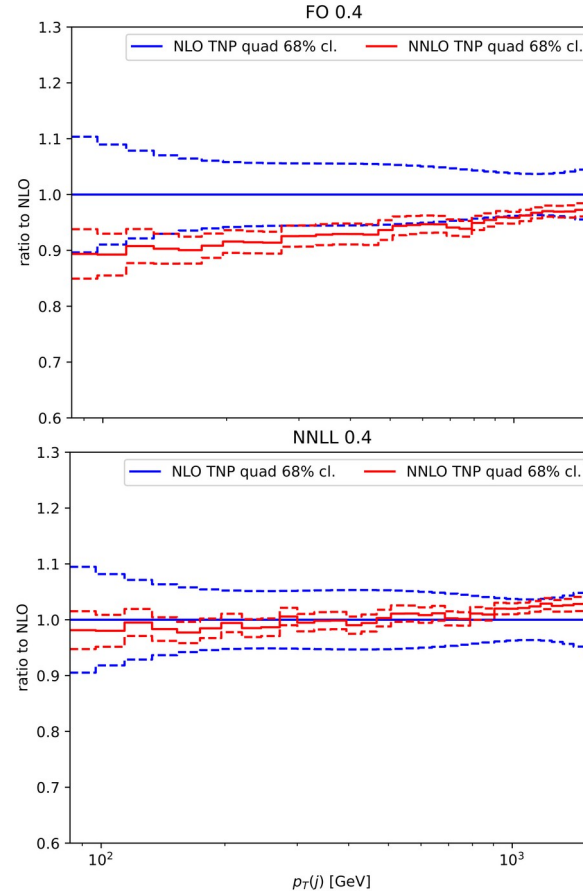


TNP uncertainties for inclusive jet production

$R = 0.4$

TNP uncertainties

- More sensible NLO uncertainties
- Similar to resummed scale variation

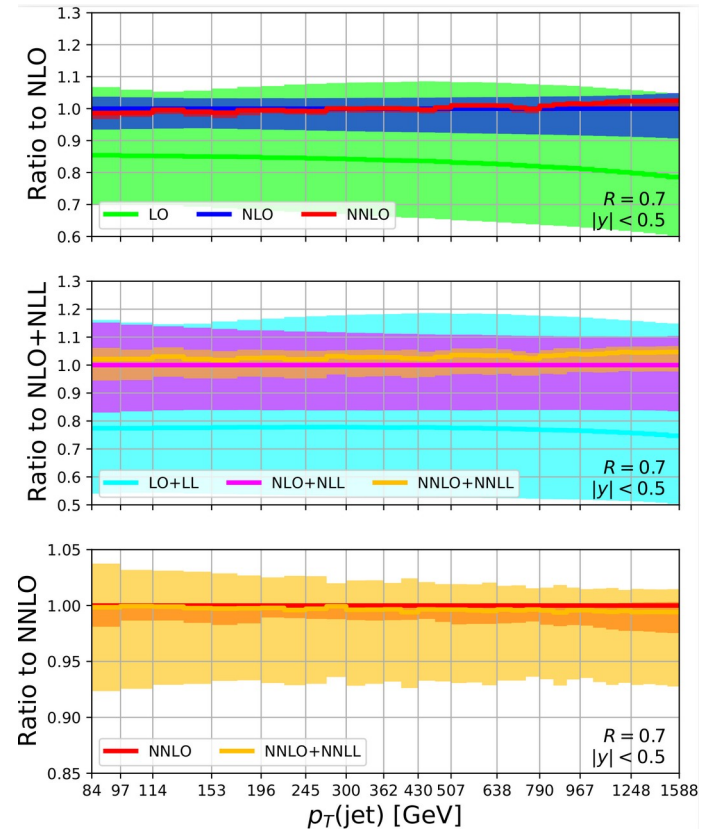
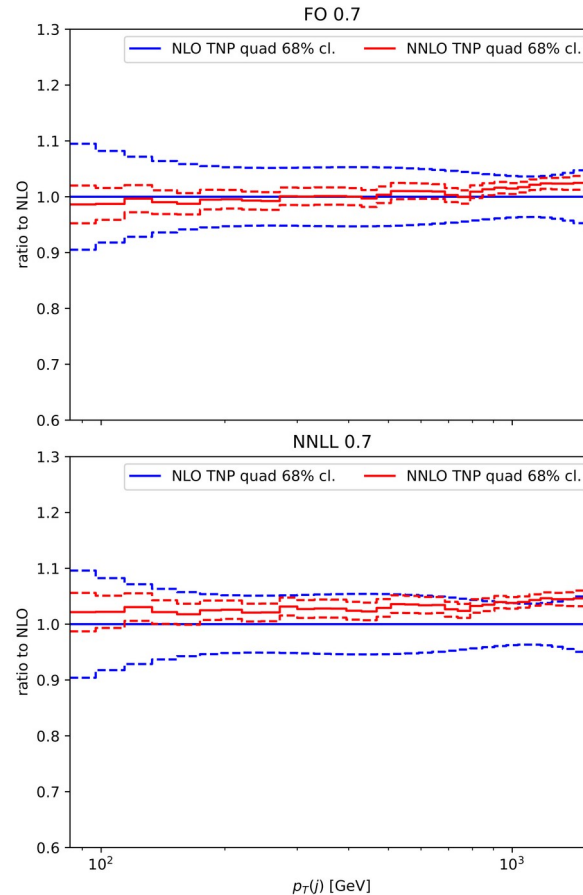


TNP uncertainties for inclusive jet production

$R = 0.7$

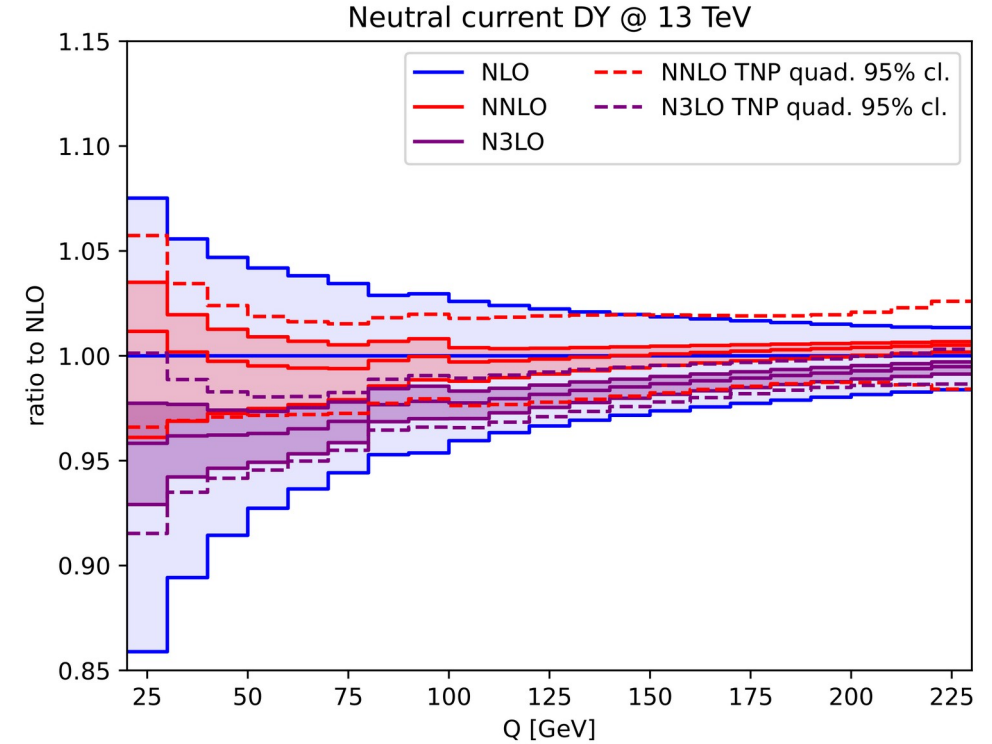
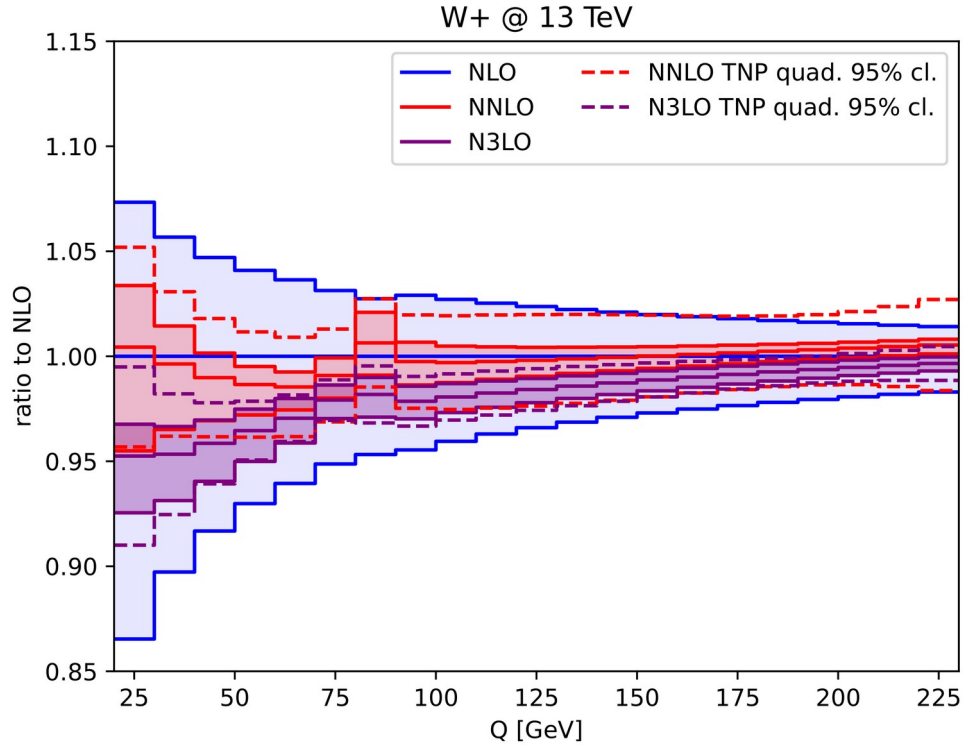
TNP uncertainties

- More sensible NNLO QCD uncertainties
- Similar to resummed scale variation



N3LO example: Drell-Yan

Numbers from n3loxs [[Baglio, Duhr, Mistlberger, Szafron'22](#)]



Points for discussion, caveats and open questions

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

- 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.**
- 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:** $d\bar{\sigma}^{(n)}(\theta) = d\Phi \langle M^0 | \mathcal{P}(\theta) | M^0 \rangle$ $\mathcal{P}(\theta)$ → **process-independent “operator”**
- How sensitive are we to the parametrisation? How many terms?
→ **two quite general parametrisations tested, increase degree by demand.**

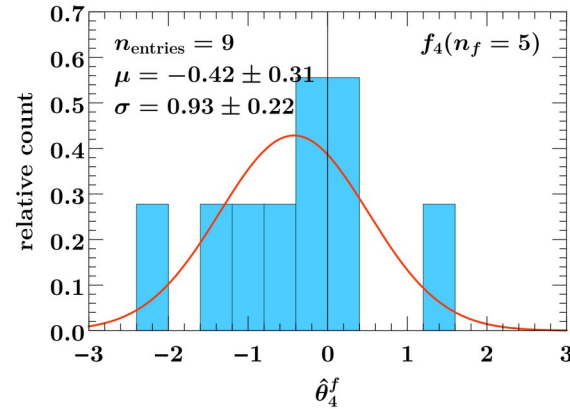
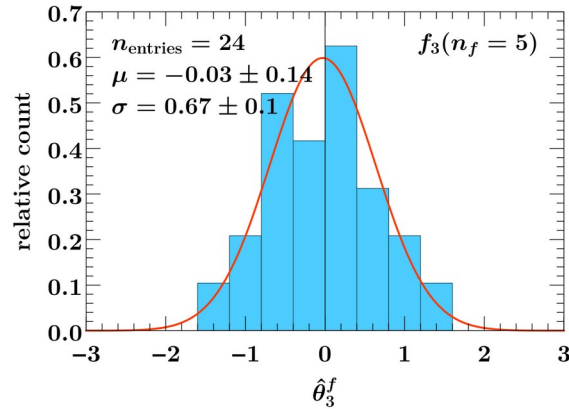
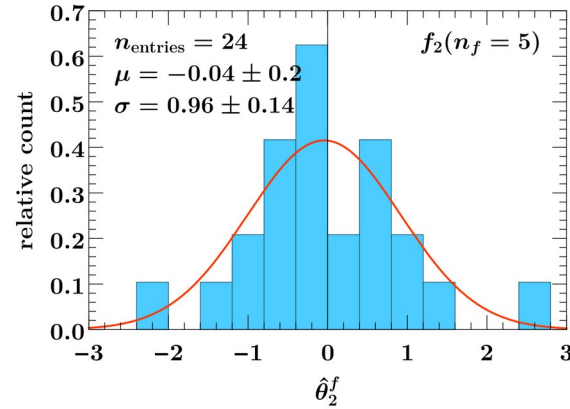
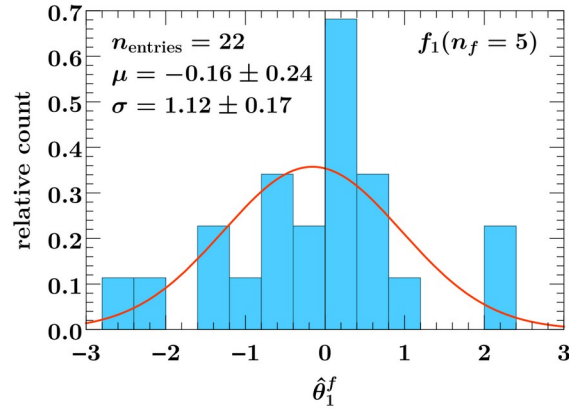
Take home message

- 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]

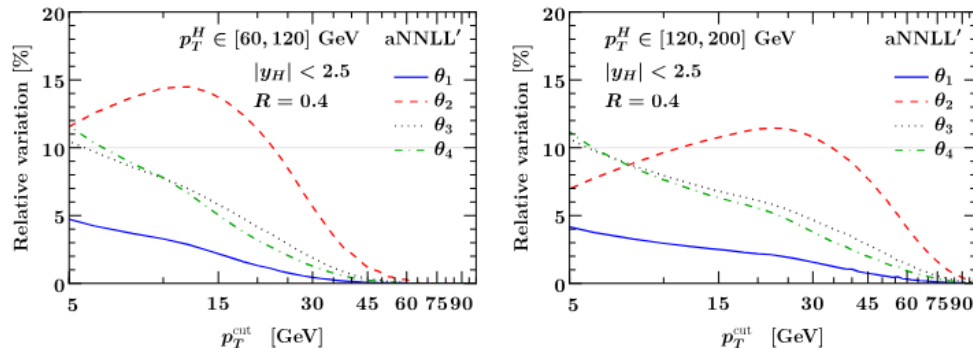
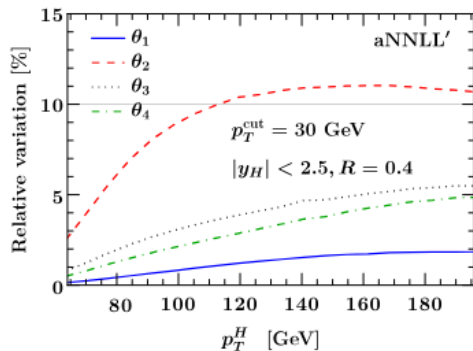


Figure 6: Relative uncertainty from varying each theory nuisance parameter as a function of p_T^{cut} for two different STXS bins.



Example: incomplete knowledge of NNLL resummation

→ some two-loop ingredients unknown

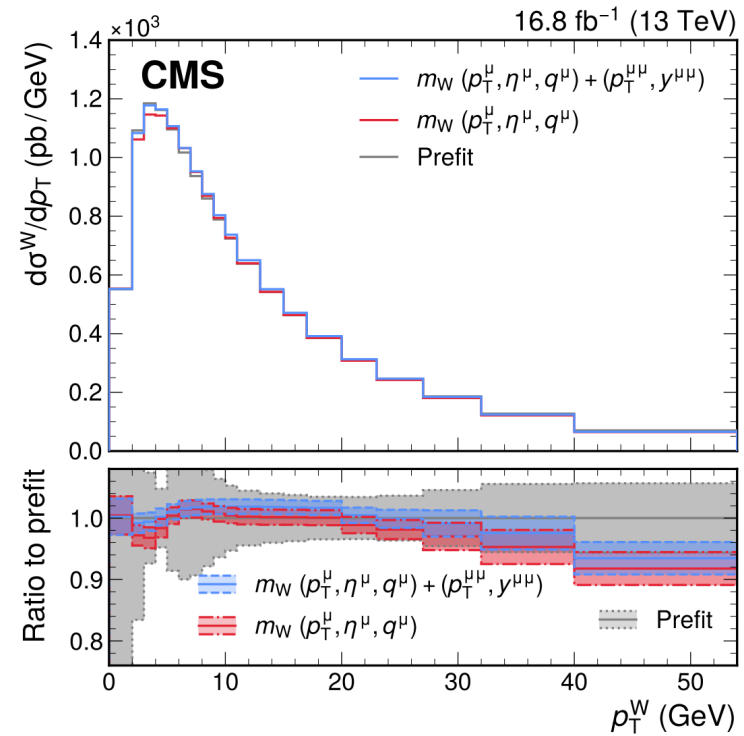
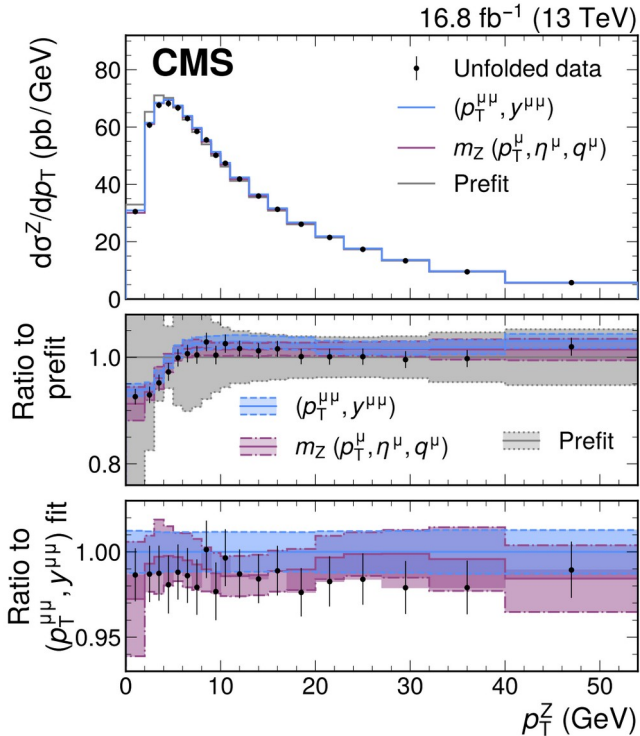
→ parametrise by TNPs

→ 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 $p_T(Z)$ → 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.

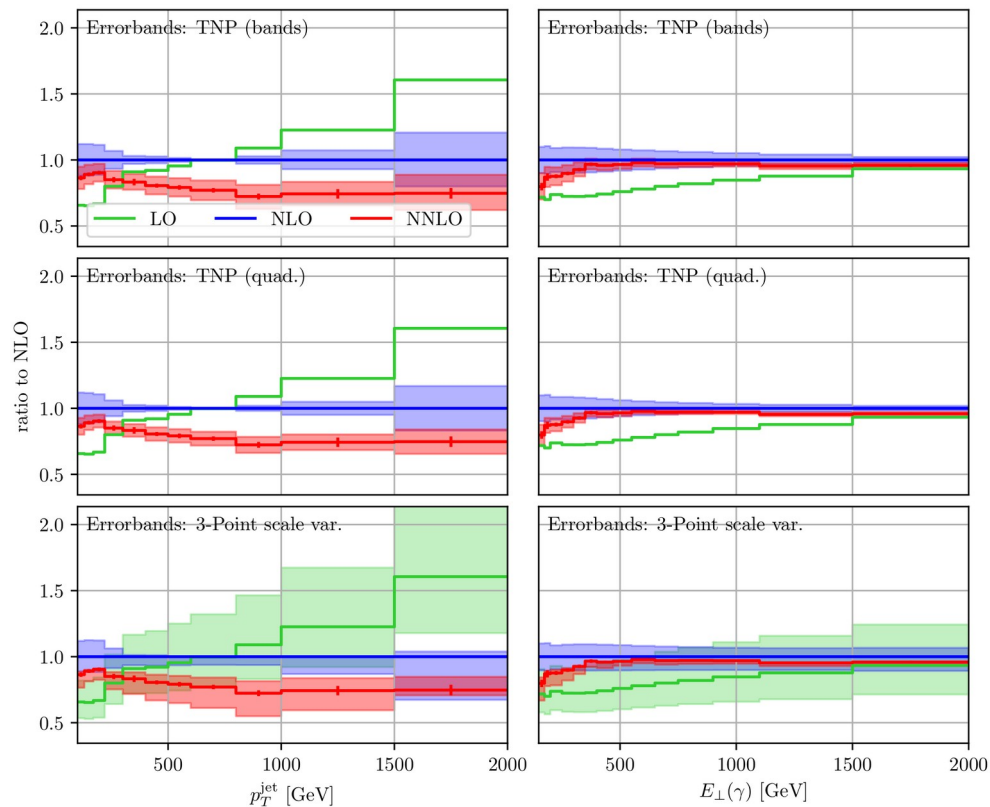
→ ideal situation

But actually not that simple:

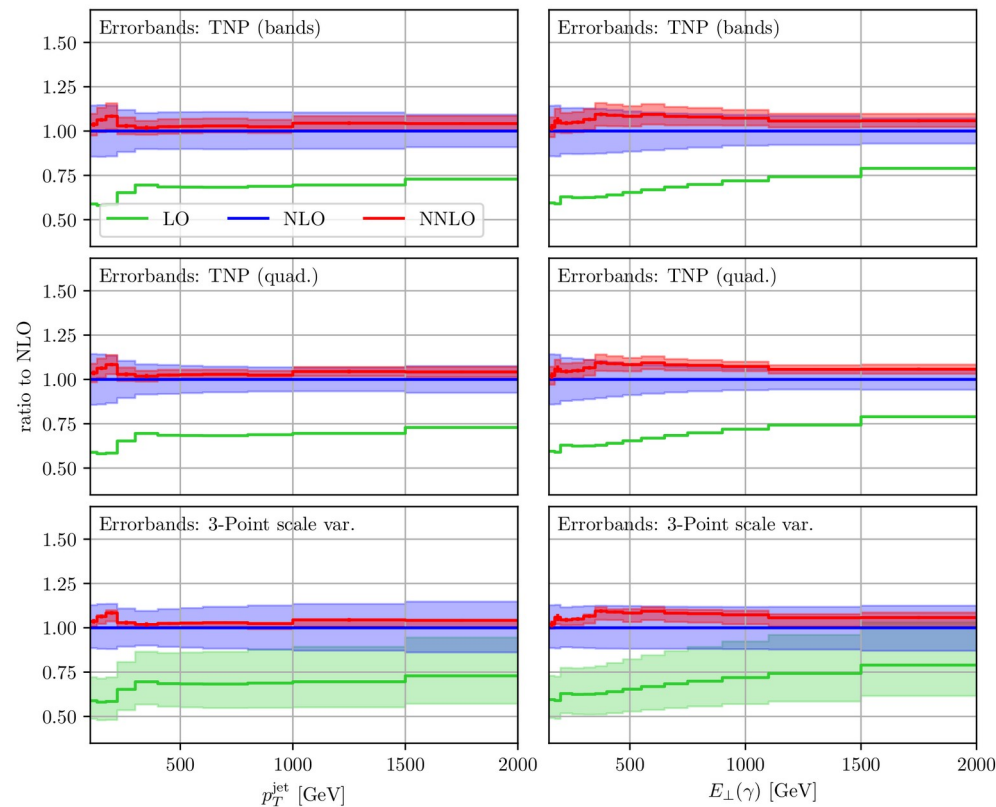
- Scheme dependence of ingredients? E.g. scales, IR subtraction, ...
→ might need modified parametrisations
- Some TNPs represent directly numbers: Γ , γ , H for simple processes
but others are functions → Beam functions, hard functions for more complicated processes
- Parametrisation works for the resummed 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

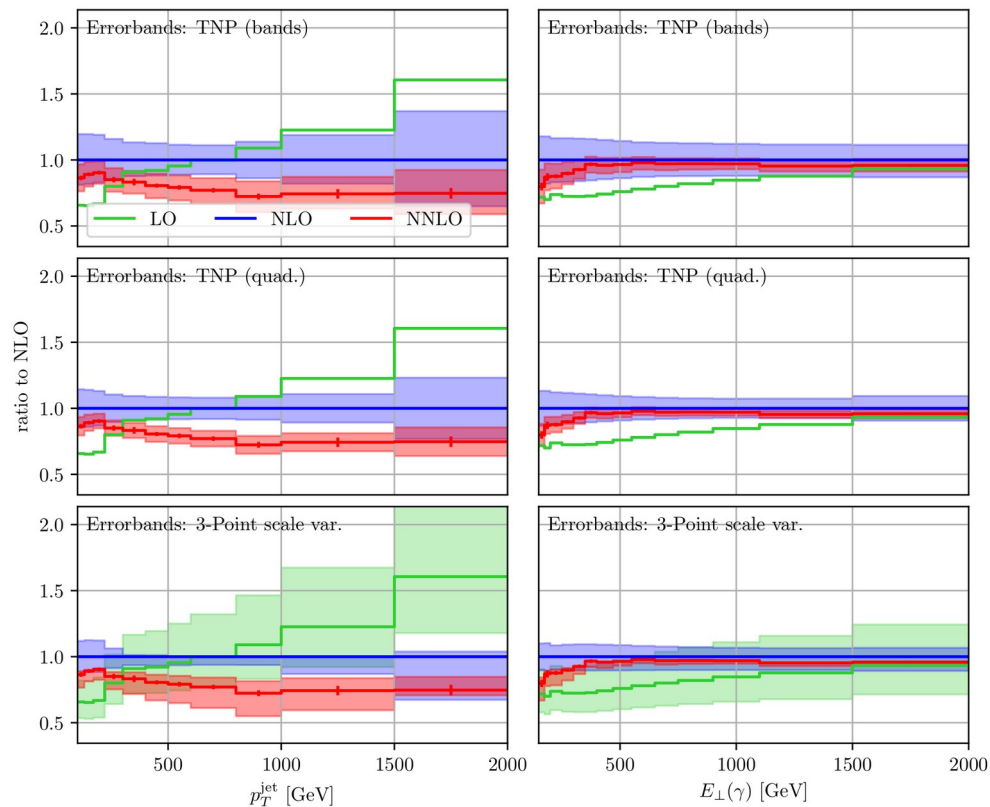


“Good” choice $\mu = H_T$ no $j=0$ term

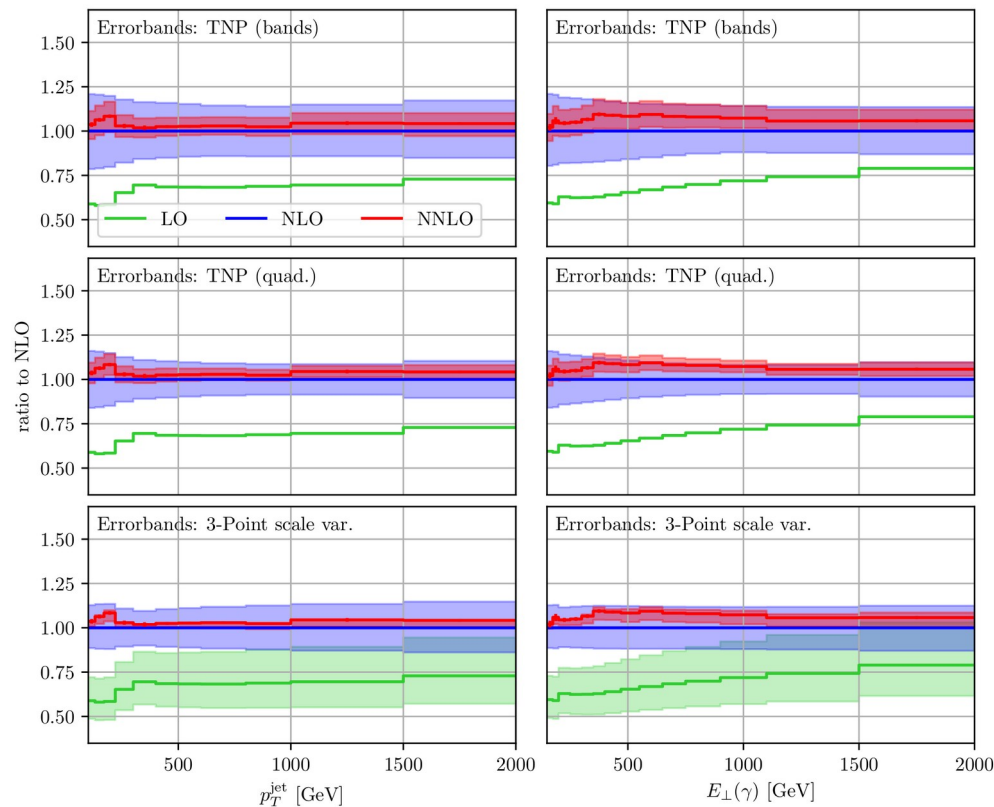


Challenging case → extended parametrisations

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

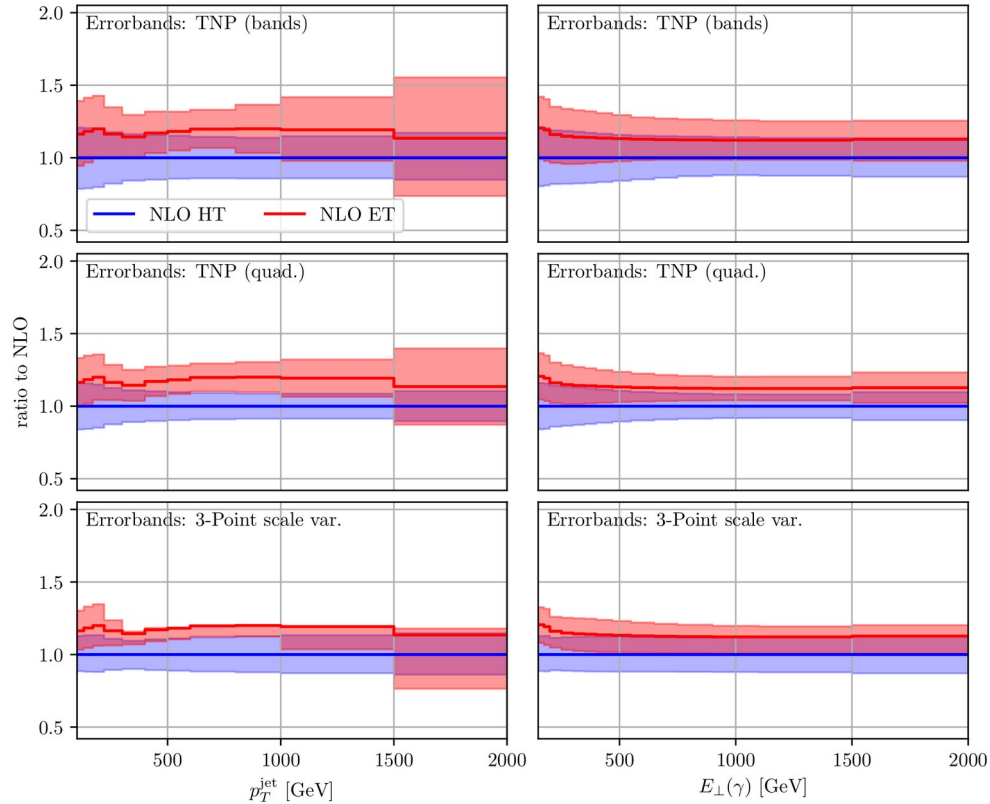


“Good” choice $\mu = H_T$ with $j=0$ term

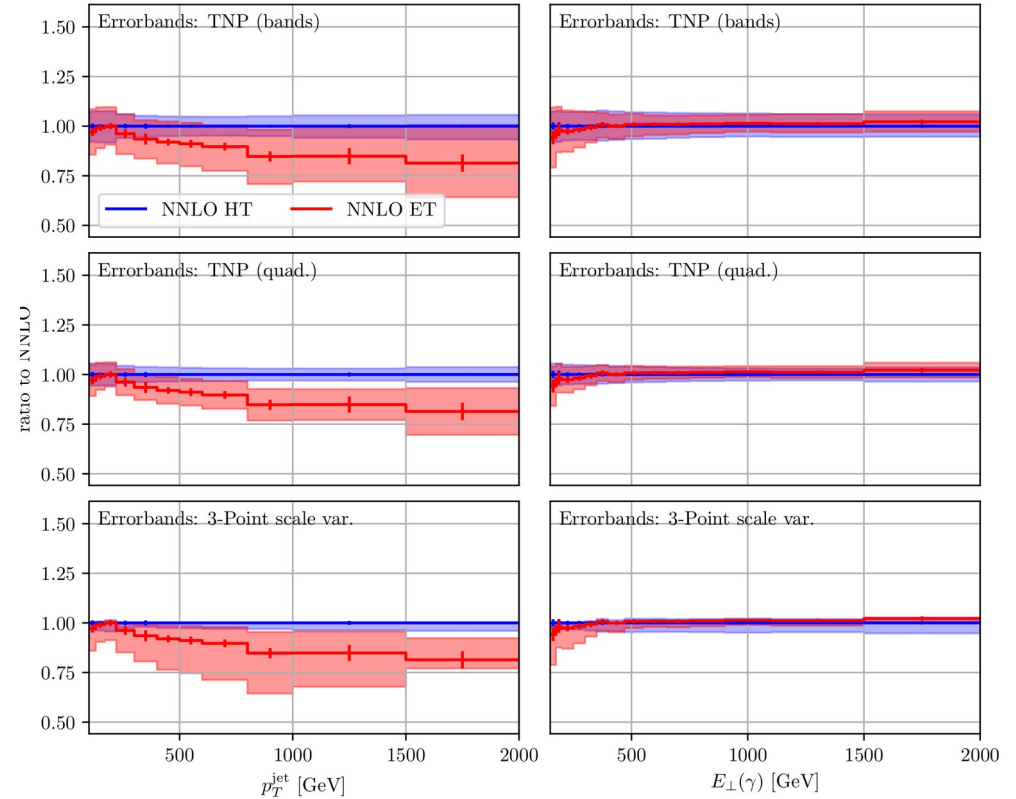


Challenging case → comparisons

NLO QCD



NNLO QCD



Bayesian approach I

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

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

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

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

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)

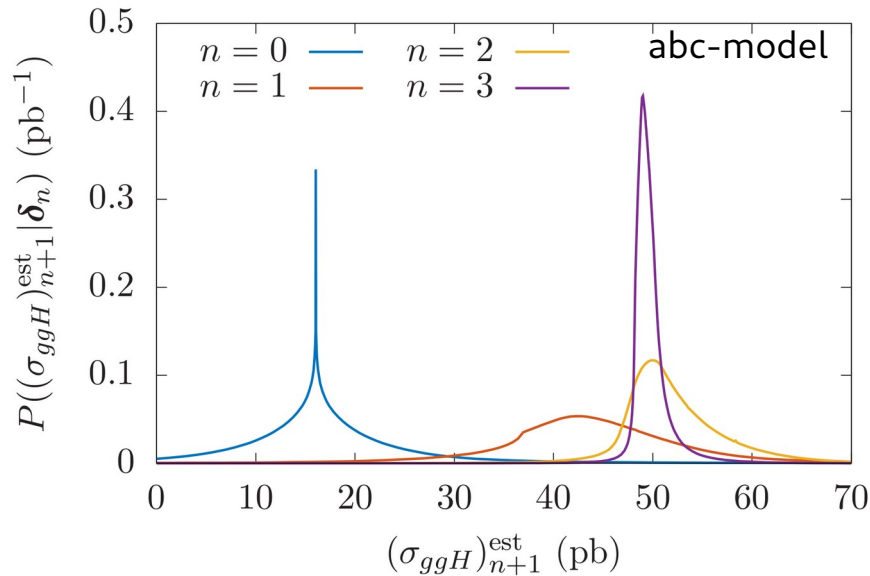
$$\mathcal{P}_{\text{sm}}(\Sigma|\Sigma_n) \approx \frac{\int d\mu P(\Sigma - \Sigma_n(\mu)|\Sigma_n(\mu)) P(\Sigma_n(\mu)) P_0(\mu)}{\int d\mu' P(\Sigma_n(\mu')) P_0(\mu')} . \quad \mu_{\text{FAC}}$$

Scale average (the results are averaged with weight function)

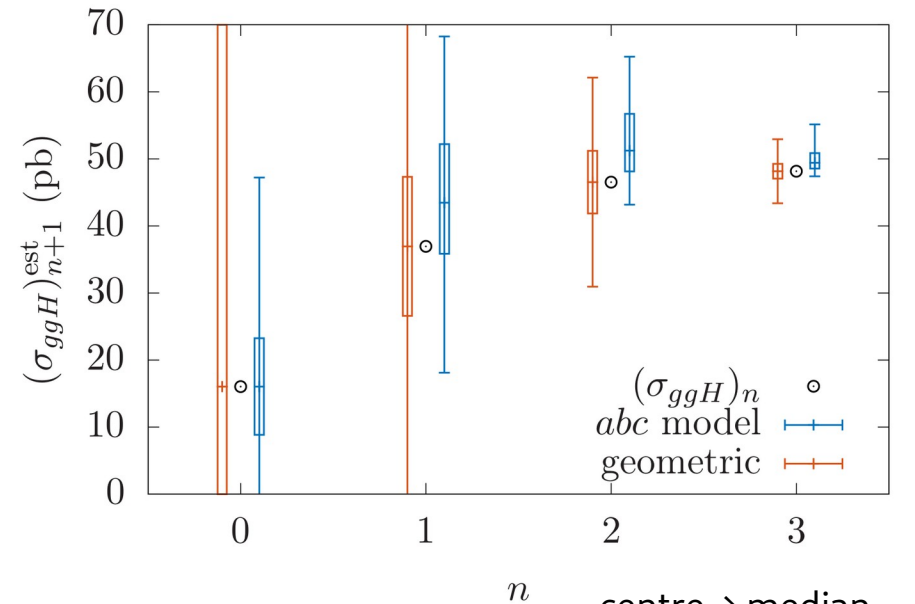
$$\mathcal{P}_{\text{sa}}(\Sigma|\Sigma_n) \approx \int d\mu w(\mu) P(\Sigma - \Sigma_n(\mu)|\Sigma_n(\mu)) \quad \mu_{\text{PMS}}$$

Bayesian approach III

Example: Higgs production in gluon - fusion



Comparison of different unc. estimates:



centre $\rightarrow$ median
error box $\rightarrow$ 68% CI
error bar $\rightarrow$ 95% CI