Perturbative Theory Uncertainties from Theory Nuisance Parameters.

Frank Tackmann

Deutsches Elektronen-Synchrotron

GDR QCD m_W Workshop
Orsay, July 1, 2025

[FT; arXiv:2411.18606]

[Thomas Cridge, Giulia Marinelli, FT; arXiv:2506.13874]

European Research Council
Established by the European Commission

Outline

Meaningful Theory Uncertainties.

Meaningful Theory Uncertainties.

Comparing a measured quantity to its theory prediction to extract POI $\boldsymbol{y}$

$$\Big[f \pm \Delta f\Big]_{\text{measured}} \overset{!}{=} \Big[f(y) \pm \Delta f\Big]_{\text{theory}} \quad \Rightarrow \quad y \pm \Delta y^{\text{exp}} \pm \Delta y^{\text{th}}$$

- *Theory uncertainty* Δf is due to inherent inexactness of our prediction because of any approximations we make

Meaningful Theory Uncertainties.

Comparing a measured quantity to its theory prediction to extract POI y

$$\Big[f \pm \Delta f\Big]_{\text{measured}} \stackrel{!}{=} \Big[f(y) \pm \Delta f\Big]_{\text{theory}} \quad \Rightarrow \quad y \pm \Delta y^{\text{exp}} \pm \Delta y^{\text{th}}$$

- *Theory uncertainty* Δf is due to inherent inexactness of our prediction because of any approximations we make

Theory uncertainty must be as meaningful as experimental one, so it should

- Reflect our level of knowledge
  - Quantify the intrinsic precision of the current prediction
  - It is *not* the distance to the true (or higher-order) result $\hat{f}$

$$\Delta f \neq |f - \hat{f}|$$

Meaningful Theory Uncertainties.

Comparing a measured quantity to its theory prediction to extract POI y

$$\Big[f \pm \Delta f\Big]_{\text{measured}} \overset{!}{=} \Big[f(y) \pm \Delta f\Big]_{\text{theory}} \quad \Rightarrow \quad y \pm \Delta y^{\text{exp}} \pm \Delta y^{\text{th}}$$

- *Theory uncertainty* Δf is due to inherent inexactness of our prediction because of any approximations we make

Theory uncertainty must be as meaningful as experimental one, so it should

- Reflect our level of knowledge
  - Quantify the intrinsic precision of the current prediction
  - It is *not* the distance to the true (or higher-order) result $\hat{f}$

$$\Delta f \neq |f - \hat{f}|$$

- Have some form of statistical interpretation
  - We want to have some probability or level of confidence that

$$\hat{f} \in [f - \Delta f, f + \Delta f]$$

Meaningful Theory Uncertainties.

Comparing a measured quantity to its theory prediction to extract POI y

$$\Big[f \pm \Delta f\Big]_{\text{measured}} \stackrel{!}{=} \Big[f(y) \pm \Delta f\Big]_{\text{theory}} \quad \Rightarrow \quad y \pm \Delta y^{\text{exp}} \pm \Delta y^{\text{th}}$$

- *Theory uncertainty* Δf is due to inherent inexactness of our prediction because of any approximations we make

Theory uncertainty must be as meaningful as experimental one, so it should

- Reflect our level of knowledge
  - Quantify the intrinsic precision of the current prediction
  - It is *not* the distance to the true (or higher-order) result $\hat{f}$

$$\Delta f \neq |f - \hat{f}|$$

- Have some form of statistical interpretation
  - We want to have some probability or level of confidence that

$$\hat{f} \in [f - \Delta f, f + \Delta f]$$

- Have correct correlations

Theory Correlations.

Correlations can be crucial once several predictions are used in combination

- Prototype example: Ratio of two quantities f and g

$$\underbrace{f}_{\text{wanted}} = \underbrace{\big[g \pm \Delta g\big]_{\text{measured}}}_{\text{measure precisely}} \times \underbrace{\left[\frac{f \pm \Delta f}{g \pm \Delta g}\right]_{\text{theory}}}_{\text{theory uncertainties cancel}}$$

  - Unc. of ratio depends critically on exact correlation ρ between Δf and Δg

Theory Correlations.

Correlations can be crucial once several predictions are used in combination

- Prototype example: Ratio of two quantities f and g

$$\underbrace{f}_{\text{wanted}} = \underbrace{\left[g \pm \Delta g\right]_{\text{measured}}}_{\text{measure precisely}} \times \underbrace{\left[\frac{f \pm \Delta f}{g \pm \Delta g}\right]_{\text{theory}}}_{\text{theory uncertainties cancel}}$$

  - Unc. of ratio depends critically on exact correlation ρ between Δf and Δg
  - For $\Delta f/f = \Delta g/g \equiv \delta$ we have $\delta_{f/g} = \delta\sqrt{2(1-\rho)}$

a correlation ρ of	99.5%	98%	95.5%	87.5%
yields ratio unc. $\delta_{f/g} =$	0.1δ	0.2δ	0.3δ	0.5δ

Theory Correlations.

Correlations can be crucial once several predictions are used in combination

- Prototype example: Ratio of two quantities f and g

$$\underbrace{f}_{\text{wanted}} = \underbrace{\left[g \pm \Delta g\right]_{\text{measured}}}_{\text{measure precisely}} \times \underbrace{\left[\frac{f \pm \Delta f}{g \pm \Delta g}\right]_{\text{theory}}}_{\text{theory uncertainties cancel}}$$

  - Unc. of ratio depends critically on exact correlation ρ between Δf and Δg
  - For $\Delta f/f = \Delta g/g \equiv \delta$ we have $\delta_{f/g} = \delta\sqrt{2(1-\rho)}$

a correlation ρ of	99.5%	98%	95.5%	87.5%
yields ratio unc. $\delta_{f/g} =$	0.1δ	0.2δ	0.3δ	0.5δ

- Differential spectrum $f(x) \equiv \mathrm{d}\sigma/\mathrm{d}x$
  - Uncertainty on *shape* of $f(x)$ is encoded in *point-by-point correlations* ρ_{ij} between $\Delta f(x_i)$ and $\Delta f(x_j)$

$\Rightarrow$ Critical for interpreting differential spectra when relying on shape effects

Scale Variations.

Our Standard Estimation Method.

Consider expansion of quantity f in some small parameter α

$$f(\alpha) = f_0 + f_1\,\alpha + f_2\,\alpha^2 + f_3\,\alpha^3 + \mathcal{O}(\alpha^4)$$

- To make a prediction, we calculate first few *true values* $\hat{f}_n$

NLO: $$f(\alpha) = \hat{f}_0 + \hat{f}_1\,\alpha$$

Our Standard Estimation Method.

Consider expansion of quantity f in some small parameter α

$$f(\alpha) = f_0 + f_1\,\alpha + f_2\,\alpha^2 + f_3\,\alpha^3 + \mathcal{O}(\alpha^4)$$

- To make a prediction, we calculate first few *true values* $\hat{f}_n$

$$\text{NLO:} \qquad f(\alpha) = \hat{f}_0 + \hat{f}_1\,\alpha$$

Can also perform the expansion in a slightly different way (aka “scheme”)

$$\tilde{\alpha}(\alpha) = \alpha\big[1 + b_0\,\alpha + b_1\,\alpha^2 + b_2\,\alpha^3 + \mathcal{O}(\alpha^4)\big]$$

$$\Rightarrow \quad \tilde{f}(\tilde{\alpha}) = \tilde{f}_0 + \tilde{f}_1\,\tilde{\alpha} + \tilde{f}_2\,\tilde{\alpha}^2 + \tilde{f}_3\,\tilde{\alpha}^3 + \mathcal{O}(\tilde{\alpha}^4)$$

- To all orders: $f(\alpha) \equiv \tilde{f}(\tilde{\alpha})$
- Finite-order prediction using $\tilde{f}(\tilde{\alpha})$ differs by higher-order terms

$$\text{NLO:} \quad \tilde{f}(\tilde{\alpha}) = \hat{\tilde{f}}_0 + \hat{\tilde{f}}_1\,\tilde{\alpha} = \hat{f}_0 + \hat{f}_1\,\alpha + b_0\,\hat{f}_1\,\alpha^2 + b_1\,\hat{f}_2\,\alpha^3 + \cdots$$

Our Standard Estimation Method.

Idea: Take the difference between schemes as uncertainty

NLO: $$\Delta f(\alpha) = b_0 \hat{f}_1\, \alpha^2 + b_1 \hat{f}_1\, \alpha^3 + \mathcal{O}(\alpha^4)$$

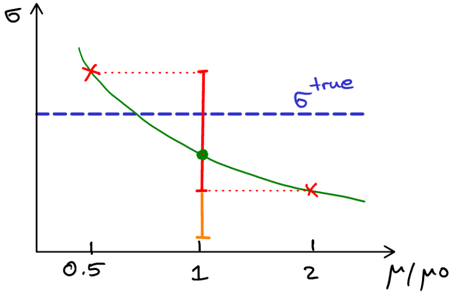
NNLO: $$\Delta f(\alpha) = \big[2 b_0 (\hat{f}_2 - b_0 \hat{f}_1) + b_1 \hat{f}_1\big] \alpha^3 + \mathcal{O}(\alpha^4)$$

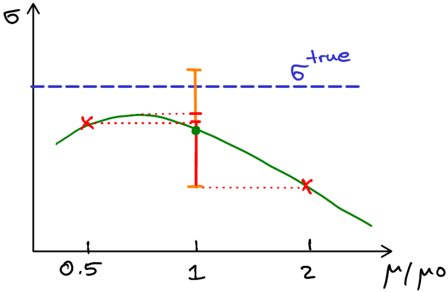
- We effectively estimate inexactness due to missing terms by approximating them by some linear combination of known lower-order terms, e.g. at NLO $f_2 \approx \hat{f}_1\, b_0$, $f_3 \approx \hat{f}_1\, b_1$

✓ Resulting estimated $\Delta f(\alpha)$ is indeed $\mathcal{O}(\alpha^{n+1})$

✗ But nothing guarantees that this is a good approximation (often it is not)
- f_{n+1} often has more complex internal structure than $f_{\leq n}$
- b_n are (rather arbitrary) numbers, have a priori nothing to do with f

Scale Variations.





$$\alpha \equiv \alpha_s(\mu_0)\,, \quad \tilde{\alpha} \equiv \alpha_s(\mu)\,, \qquad b_0 = 0.85\,L\,, \quad b_1 = 0.75\,L^2 + 0.34\,L$$

Correspond to a continuous scheme change with $L = \ln(\mu/\mu_0)/\ln 2$

- μ (or b_0) is *not* an actual parameter with a true value that f depends on
  - No value for it might ever capture the true result (happens regularly)
  - At higher order, uncertainty reduces *only because* μ (or b_0) becomes less relevant and *not because* it somehow becomes better known

What About Correlations?

$$f(\alpha) = \hat{f}_0 + \hat{f}_1\,\alpha \pm \Delta f \qquad \text{with} \qquad \Delta f = \hat{f}_1\, b_0\, \alpha^2 + \cdots$$
$$g(\alpha) = \hat{g}_0 + \hat{g}_1\,\alpha \pm \Delta g \qquad \text{with} \qquad \Delta g = g_1\, b_0\, \alpha^2 + \cdots$$

How are Δf and Δg correlated?

- We don't know – the method simply does not tell us
  - Correlations require a common uncertain parameter
    (or more generally a common source of uncertainty)
  - $b_0, b_1, \ldots$ (or μ) are not common or uncertain parameters
  - ✗ Scale variations *do not* give correct correlations or shape uncertainties
- Best we can do is *assume* some theoretically motivated but still *ad hoc* correlation model that we impose on Δf and Δg
  - Example: Correlating/uncorrelating scale variations
  - Another example: Offsetting/scanning over scale variations (→ backup)

⇒ Probably the most severe shortcoming of scale variations

Existing Alternatives to Scale Variations.

There have been precious few efforts to develop alternatives

- Bayesian models [Cacciari, Houdeau '11; Bagnaschi et al. '14; Bonvini '20; Duhr et al. '21]
- Series acceleration [David, Passarino '13]
- Using reference processes [Gosh et al. '22]
- All go in the right direction
  - ✓ Try to more directly estimate size of missing higher orders
  - ✓ Expose assumptions more explicitly
  - ✓ Try to address statistical interpretation

However

- ✗ Uncertainty estimate still based on the known lower-order terms and/or scale variations
- ✗ Do not address correlations

⇒ Unfortunately, have similar level of arbitrariness and share many limitations of scale variations

Theory Nuisance Parameters.

Parametric Theory Uncertainties.

Let's go back to our expansion of f

$$f(\alpha) = f_0 + f_1\,\alpha + f_2\,\alpha^2 + f_3\,\alpha^3 + \mathcal{O}(\alpha^4)$$

- We calculate first few true $\hat{f}_n$ to make an approximate prediction

$$\begin{aligned} \text{LO:} \quad & f(\alpha) = \hat{f}_0 \\ \text{NLO:} \quad & f(\alpha) = \hat{f}_0 + \hat{f}_1\,\alpha \\ \text{NNLO:} \quad & f(\alpha) = \hat{f}_0 + \hat{f}_1\,\alpha + \hat{f}_2\,\alpha^2 \end{aligned}$$

- Q: What is the source of uncertainty in this?
  - The missing terms $f_n\,\alpha_s^n$?
  - No, actually, it is the unknown series coefficients f_n
  - Or more precisely, their unknown true values $\hat{f}_n$

⇒ To estimate the theory uncertainty, we need to quantify our limited or lack of knowledge of f_n

Parametric Theory Uncertainties.

Consider f_n as an unknown parameter to be varied in some way

- To evaluate associated uncertainty, we need to propagate this variation into the prediction
- For that, f_n actually has to appear in the prediction, so we include it

$$\text{N}^{1+1}\text{LO:} \qquad f(\alpha, f_2) = \hat{f}_0 + \hat{f}_1\,\alpha + f_2\,\alpha^2$$

$$\text{N}^{1+2}\text{LO:} \qquad f(\alpha, f_2, f_3) = \hat{f}_0 + \hat{f}_1\,\alpha + f_2\,\alpha^2 + f_3\,\alpha^3$$

$$\text{N}^{2+1}\text{LO:} \qquad f(\alpha, f_3) = \hat{f}_0 + \hat{f}_1\,\alpha + \hat{f}_2\,\alpha^2 + f_3\,\alpha^3$$

- And that's basically it!
  - We explicitly include the (leading) sources of uncertainty in our prediction
  - f_n are well-defined parameters of our prediction with a true but unknown (or imprecisely known) value
  - We simply treat them as such

Theory Nuisance Parameters (TNPs).

In reality, $f_n \equiv f_n(x)$ has (lots of) nontrivial internal structure "x"

- Parameterize it in terms of TNPs: $f_n(x, \theta_n)$ where $\theta_n \equiv \{\theta_n^i\}$

N^{1+1}LO: $f(\alpha, x, \theta_2) = \hat{f}_0(x) + \hat{f}_1(x)\alpha + f_2(x, \theta_2)\alpha^2$

N^{1+2}LO: $f(\alpha, x, \theta_{2,3}) = \hat{f}_0(x) + \hat{f}_1(x)\alpha + f_2(x, \theta_2)\alpha^2 + f_3(x, \theta_3)\alpha^3$

N^{2+1}LO: $f(\alpha, x, \theta_3) = \hat{f}_0(x) + \hat{f}_1(x)\alpha + \hat{f}_2(x)\alpha^2 + f_3(x, \theta_3)\alpha^3$

Theory Nuisance Parameters (TNPs).

In reality, $f_n \equiv f_n(x)$ has (lots of) nontrivial internal structure "x"

- Parameterize it in terms of TNPs: $f_n(x, \theta_n)$ where $\theta_n \equiv \{\theta_n^i\}$

N^{1+1}LO: $f(\alpha, x, \theta_2) = \hat{f}_0(x) + \hat{f}_1(x)\alpha + f_2(x, \theta_2)\alpha^2$

N^{1+2}LO: $f(\alpha, x, \theta_{2,3}) = \hat{f}_0(x) + \hat{f}_1(x)\alpha + f_2(x, \theta_2)\alpha^2 + f_3(x, \theta_3)\alpha^3$

N^{2+1}LO: $f(\alpha, x, \theta_3) = \hat{f}_0(x) + \hat{f}_1(x)\alpha + \hat{f}_2(x)\alpha^2 + f_3(x, \theta_3)\alpha^3$

- Key condition: There must be true values $\hat{\theta}_n$ such that

$$\hat{f}_n(x) = f_n(x, \hat{\theta}_n)$$

  - This makes TNPs well-defined parameters with true but unknown value
  - Theory uncertainties become truly *parametric* with all the implied benefits

Theory Nuisance Parameters (TNPs).

In reality, $f_n \equiv f_n(x)$ has (lots of) nontrivial internal structure "x"

- Parameterize it in terms of TNPs: $f_n(x, \theta_n)$ where $\theta_n \equiv \{\theta_n^i\}$

N^{1+1}LO: $f(\alpha, x, \theta_2) = \hat{f}_0(x) + \hat{f}_1(x)\alpha + f_2(x, \theta_2)\alpha^2$

N^{1+2}LO: $f(\alpha, x, \theta_{2,3}) = \hat{f}_0(x) + \hat{f}_1(x)\alpha + f_2(x, \theta_2)\alpha^2 + f_3(x, \theta_3)\alpha^3$

N^{2+1}LO: $f(\alpha, x, \theta_3) = \hat{f}_0(x) + \hat{f}_1(x)\alpha + \hat{f}_2(x)\alpha^2 + f_3(x, \theta_3)\alpha^3$

- Key condition: There must be true values $\hat{\theta}_n$ such that

$$\hat{f}_n(x) = f_n(x, \hat{\theta}_n)$$

  - This makes TNPs well-defined parameters with true but unknown value
  - Theory uncertainties become truly *parametric* with all the implied benefits

⇒ Step 1: Derive an appropriate TNP parameterization $f_n(x, \theta_n)$
and implement it in all the predictions

Step 1: TNP Parameterization.

$f_n(x)$ are functions of various x dependences

- Discrete: partonic channels, quantum numbers, ...
- Continuous but discrete values: n_f, N_c, $E_{\rm cm}$, ...
- Fully continuous: Kinematic variables (p_T^Z, Y, q^2), particle masses, ...

- Q: Which dependences do we need to account for and parameterize?
  - All those in which we need correlations
  - Plus those that help to obtain better theory constraints

Step 1: TNP Parameterization.

$f_n(x)$ are functions of various x dependences

- Discrete: partonic channels, quantum numbers, ...
- Continuous but discrete values: n_f, N_c, E_{cm}, ...
- Fully continuous: Kinematic variables (p_T^Z, Y, q^2), particle masses, ...

- Q: Which dependences do we need to account for and parameterize?
  - All those in which we need correlations
  - Plus those that help to obtain better theory constraints

To satisfy key condition $f_n(x, \theta_n)$ must encode correct x dependence

$$\hat{f}_n(x) = f_n(x, \hat{\theta}_n)$$

- Requires expert knowledge on underlying structure (functional form in x)
  - Deriving $f_n(x, \theta_n)$ means deriving correct theory correlation structure in x

✗ This is where scale variations, Pythia variations, etc. fail:
  - Do not yield a correct parameterization, cannot give correct correlations

Step 1: Parameterization Strategies.

General strategy

⇒ Break down internal structure until remaining unknowns $f_{n,i}$ are numbers

Specific strategy depends on what we know about functional form of $f_n(x)$

1. We know it well enough to parameterize it explicitly

   For example:

   - Dependence on partonic channels is always known exactly
   - We might know from theory that $f_n(x)$ is a polynomial in $\ln x$

$$\Rightarrow \qquad f_n(x) = \sum_{i=0}^{k} f_{n,i} \ln^i x$$

Step 1: Parameterization Strategies.

General strategy

⇒ Break down internal structure until remaining unknowns $f_{n,i}$ are numbers

Specific strategy depends on what we know about functional form of $f_n(x)$

1. We know it well enough to parameterize it explicitly
2. We know it well enough to apply 1) in some limit and can expand around that limit

$$f_n(x) = f_{n0}(x) + f_{n1}(x)\,\varepsilon + f_{n2}(x)\,\varepsilon^2 + \mathcal{O}(\varepsilon^3)$$

For example:

- Drell-Yan p_T spectrum at small p_T: $\varepsilon = p_T^2/Q^2$
- $f_{n0}(p_T)$ can be obtained from p_T resummation with strategy 1)

Step 1: Parameterization Strategies.

General strategy

⇒ Break down internal structure until remaining unknowns $f_{n,i}$ are numbers

Specific strategy depends on what we know about functional form of $f_n(x)$

1. We know it well enough to parameterize it explicitly
2. We know it well enough to apply 1) in some limit
3. If we don't have sufficient knowledge for either strategy 1 or 2, we can always expand in a suitable, complete functional basis $\{\phi_i\}$

$$f_n(x) = \sum_{i=0}^{\infty} f_{n,i}\, \phi_i(x) \quad \to \quad f_n(x, \theta_n) = \sum_{i=0}^{k} \theta_{n,i}\, \phi_i(x)$$

- A good basis is one that we can truncate after a few terms with high confidence that we can neglect truncated terms as a subleading source of uncertainty
- There are various ways for designing good bases

Constraining the TNPs.

Since θ_n are proper parameters, they can have proper estimates/constraints

$$\theta_n = u_n \pm \Delta u_n$$

- Can come from
  - From theory: Can be modelled as “imagined” auxiliary measurement
  - From experiment: Real auxiliary or nominal measurement

Constraining the TNPs.

Since θ_n are proper parameters, they can have proper estimates/constraints

$$\theta_n = u_n \pm \Delta u_n$$

- Can come from
  - From theory: Can be modelled as “imagined” auxiliary measurement
  - From experiment: Real auxiliary or nominal measurement
- Could profile them as a free nuisance parameter in fit to data
  - No dependence on theory prejudice on how much to vary/assumptions
  - If the data is sensitive to a θ_n, it will constrain it, otherwise it doesn’t matter

Constraining the TNPs.

Since θ_n are proper parameters, they can have proper estimates/constraints

$$\theta_n = u_n \pm \Delta u_n$$

- Can come from
  - From theory: Can be modelled as “imagined” auxiliary measurement
  - From experiment: Real auxiliary or nominal measurement
- Could profile them as a free nuisance parameter in fit to data
  - No dependence on theory prejudice on how much to vary/assumptions
  - If the data is sensitive to a θ_n, it will constrain it, otherwise it doesn’t matter
- Nevertheless, still worthwhile/useful to have a theory constraint
  - Leaving them completely free in a fit may not always be possible or practical
  - Important to know their expected “natural” size to check data constraints
  - We do like to know the uncertainty of a prediction without having to fit to data

$\Rightarrow$ Step 2: Obtain suitable constraints on θ_n

Step 2: Theory Constraints for Scalar Series.

Assume everything has been broken down to scalar series (in QCD)

$$f(\alpha_s) = 1 + \sum_{n=1} f_n \Big(\frac{\alpha_s}{4\pi}\Big)^n$$

- Parameterize coefficients f_n (numbers) as

$$f_n(\theta_n) = N_n \, \theta_n$$

  - Normalization N_n from expected "natural size" of f_n: $|\hat{f}_n| \lesssim N_n$
  - Expect θ_n to have $\mathcal{O}(1)$ natural size: $|\hat{\theta}_n| \lesssim 1$

- Impose a theory constraint

$$\theta_n = u_n \pm \Delta u_n = 0 \pm 1$$

  - Default assumption: Model u_n as normal distributed random variable

$\Rightarrow$ Defines a specific "theory estimator": Validate it on known pert. series

Step 2: Estimating Natural Size.

Let's take as an example: $q\bar{q} \to V$, $q\bar{q} \to H$, $gg \to H$ form factors

$f(\alpha_s)$	N_n	$\hat{f}_1/N_1$	$\hat{f}_2/N_2$	$\hat{f}_3/N_3$	$\hat{f}_4/N_4$
$c_{q\bar{q}V}(\alpha_s)$	1	-8.47	-48.6	-1387	-42015
$c_{q\bar{q}S}(\alpha_s)$	1	-0.47	$+87.1$	$+2309$	$+76100$
$c_{gg}(\alpha_s)$	1	$+4.93$	-24.0	-4066	-123979

Step 2: Estimating Natural Size.

Let's take as an example: $q\bar{q} \to V$, $q\bar{q} \to H$, $gg \to H$ form factors

$f(\alpha_s)$	N_n	$\hat{f}_1/N_1$	$\hat{f}_2/N_2$	$\hat{f}_3/N_3$	$\hat{f}_4/N_4$
$c_{q\bar{q}V}(\alpha_s)$	1	-8.47	-48.6	-1387	-42015
	4^n	-2.12	-3.04	-21.7	-164
$c_{q\bar{q}S}(\alpha_s)$	1	-0.47	$+87.1$	$+2309$	$+76100$
	4^n	-0.12	$+5.44$	$+36.1$	$+297$
$c_{gg}(\alpha_s)$	1	$+4.93$	-24.0	-4066	-123979
	4^n	$+1.23$	-1.50	-63.5	-484

Step 2: Estimating Natural Size.

Let's take as an example: $q\bar{q} \to V$, $q\bar{q} \to H$, $gg \to H$ form factors

$f(\alpha_s)$	N_n	$\hat{f}_1/N_1$	$\hat{f}_2/N_2$	$\hat{f}_3/N_3$	$\hat{f}_4/N_4$
$c_{q\bar{q}V}(\alpha_s)$	1	-8.47	-48.6	-1387	-42015
	4^n	-2.12	-3.04	-21.7	-164
	$4^n C_F C_A^{n-1}$	-1.59	-0.76	-1.81	-4.56
$c_{q\bar{q}S}(\alpha_s)$	1	-0.47	$+87.1$	$+2309$	$+76100$
	4^n	-0.12	$+5.44$	$+36.1$	$+297$
	$4^n C_F C_A^{n-1}$	-0.09	$+1.36$	$+3.01$	$+8.26$
$c_{gg}(\alpha_s)$	1	$+4.93$	-24.0	-4066	-123979
	4^n	$+1.23$	-1.50	-63.5	-484
	$4^n C_A C_A^{n-1}$	$+0.41$	-0.17	-2.35	-5.98

Step 2: Estimating Natural Size.

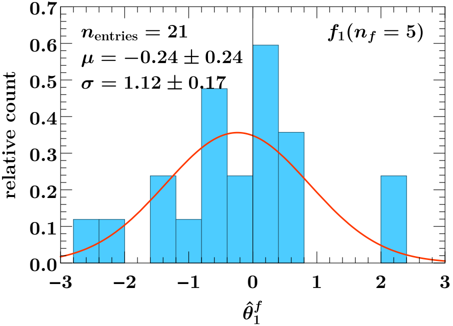
Let's take as an example: $q\bar{q} \to V$, $q\bar{q} \to H$, $gg \to H$ form factors

$f(\alpha_s)$	N_n	$\hat{f}_1/N_1$	$\hat{f}_2/N_2$	$\hat{f}_3/N_3$	$\hat{f}_4/N_4$
$c_{q\bar{q}V}(\alpha_s)$	1	-8.47	-48.6	-1387	-42015
	4^n	-2.12	-3.04	-21.7	-164
	$4^n C_F C_A^{n-1}$	-1.59	-0.76	-1.81	-4.56
	$4^n C_F C_A^{n-1}(n-1)!$	-1.59	-0.76	-0.90	-0.76
$c_{q\bar{q}S}(\alpha_s)$	1	-0.47	$+87.1$	$+2309$	$+76100$
	4^n	-0.12	$+5.44$	$+36.1$	$+297$
	$4^n C_F C_A^{n-1}$	-0.09	$+1.36$	$+3.01$	$+8.26$
	$4^n C_F C_A^{n-1}(n-1)!$	-0.09	$+1.36$	$+1.50$	$+1.38$
$c_{gg}(\alpha_s)$	1	$+4.93$	-24.0	-4066	-123979
	4^n	$+1.23$	-1.50	-63.5	-484
	$4^n C_A C_A^{n-1}$	$+0.41$	-0.17	-2.35	-5.98
	$4^n C_A C_A^{n-1}(n-1)!$	$+0.41$	-0.17	-1.18	-1.00

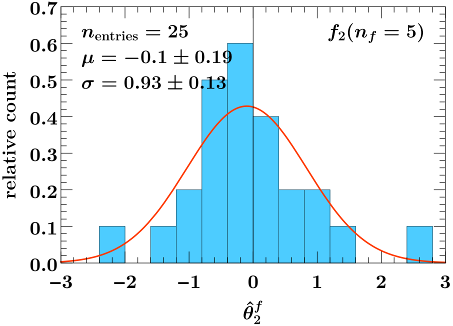
$\Rightarrow$ Let's pick: $N_n = 4C_r(4C_A)^{n-1}(n-1)!$

Step 2: Validation on Known Series.

1 loop: $\hat{\theta}_1^f$



2 loop: $\hat{\theta}_2^f$

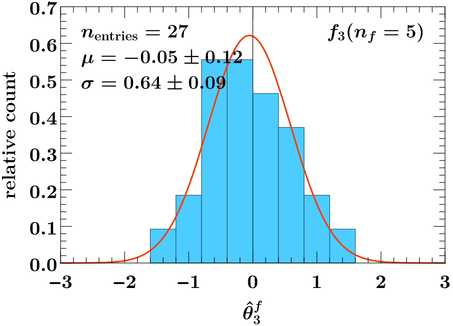


Consider set of QCD series $F = \{f\}$ of a common type/category

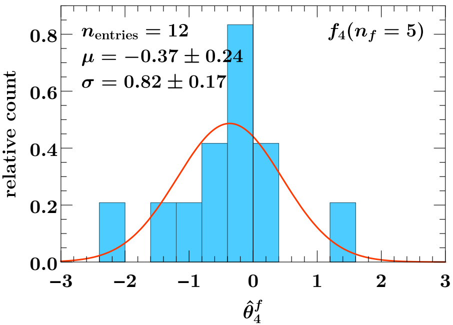
- Identify $P_{\theta_n^f}(x)$ with *population distribution* of $\hat{\theta}_n^f \in F_n$ (assumption)
  - For any given θ_n^f, its unknown $\hat{\theta}_n^f$ comes from a QCD bag of coefficients
- Estimate population distribution from a sample of known series
  - Good fit to normal distribution ("central-limit theorem of Feynman diagrams")
  - Note: Distribution is a property of this specific estimator

Step 2: Validation on Known Series.

3 loop: $\hat{\theta}_3^f$



4 loop: $\hat{\theta}_4^f$

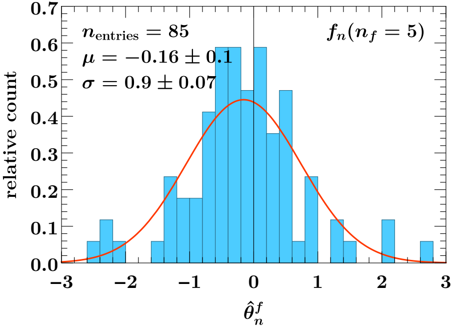


Consider set of QCD series $F = \{f\}$ of a common type/category

- Identify $P_{\theta_n^f}(x)$ with *population distribution* of $\hat{\theta}_n^f \in F_n$ (assumption)
  - For any given θ_n^f, its unknown $\hat{\theta}_n^f$ comes from a QCD bag of coefficients
- Estimate population distribution from a sample of known series
  - Good fit to normal distribution (“central-limit theorem of Feynman diagrams”)
  - Note: Distribution is a property of this specific estimator

Step 2: Validation on Known Series.

1-4 loop: $\hat{\theta}_n^f$



Consider set of QCD series $F = \{f\}$ of a common type/category

- Identify $P_{\theta_n^f}(x)$ with *population distribution* of $\hat{\theta}_n^f \in F_n$ (assumption)
  - For any given θ_n^f, its unknown $\hat{\theta}_n^f$ comes from a QCD bag of coefficients
- Estimate population distribution from a sample of known series
  - Good fit to normal distribution (“central-limit theorem of Feynman diagrams”)
  - Note: Distribution is a property of this specific estimator

Application to Drell-Yan p_T Spectrum.

TNP Parameterization for V p_T Spectrum.

Consider dependence on $x \equiv q_T = p_T^V$ (where $V = Z/\gamma, W$)

1) Apply strategy 2: Expand in $\varepsilon = q_T^2/Q^2$ (where $Q \equiv m_{\ell\ell}$)

$$\frac{d\sigma}{dq_T} = \frac{d\sigma^{(0)}}{dq_T} \times \left[1 + \mathcal{O}\left(\frac{q_T^2}{Q^2}\right)\right]$$

- $\mathcal{O}(q_T^2/Q^2)$ corrections stay below $\lesssim 5\%$ up to $q_T \lesssim Q/3 \ldots Q/2$

TNP Parameterization for V p_T Spectrum.

Consider dependence on $x \equiv q_T = p_T^V$ (where $V = Z/\gamma, W$)

1) Apply strategy 2: Expand in $\varepsilon = q_T^2/Q^2$ (where $Q \equiv m_{\ell\ell}$)

$$\frac{\mathrm{d}\sigma}{\mathrm{d}q_T} = \frac{\mathrm{d}\sigma^{(0)}}{\mathrm{d}q_T} \times \left[1 + \mathcal{O}\left(\frac{q_T^2}{Q^2}\right)\right]$$

- $\mathcal{O}(q_T^2/Q^2)$ corrections stay below $\lesssim 5\%$ up to $q_T \lesssim Q/3 \ldots Q/2$

2) Apply strategy 1 to $\mathrm{d}\sigma^{(0)}/\mathrm{d}q_T$

$$\frac{\mathrm{d}\sigma^{(0)}}{\mathrm{d}q_T} = \left[\sum_{a,b} H_{ab} \times B_a \otimes B_b \otimes S\right](L = \ln q_T/Q)$$

$$F(\alpha_s, L) = F(\alpha_s) \exp \int_0^L \mathrm{d}L' \left\{\Gamma_{\mathrm{cusp}}[\alpha_s(L')]\, L' + \gamma_F[\alpha_s(L')]\right\}$$

- q_T dependence is predicted by resummation in terms of several independent (scalar) series
- Boundary conditions and anomalous dimensions of RGE for each function

TNPs for p_T Spectrum: After the Dust has Settled.

Brake things down to independent perturbative series, e.g. at N^{2+1}LL

- 5 scalar series (plus a few more we can neglect here for simplicity)

$$\Gamma(\alpha_s) = \alpha_s\,\hat\Gamma_0 + \alpha_s^2\,\hat\Gamma_1 + \alpha_s^3\,\hat\Gamma_2 + \alpha_s^4\,\Gamma_3(\theta_3^\Gamma)$$

$$\gamma_\mu(\alpha_s) = \alpha_s\,\hat\gamma_{\mu 0} + \alpha_s^2\,\hat\gamma_{\mu 1} + \alpha_s^3\,\gamma_{\mu 2}(\theta_2^{\gamma_\mu})$$

$$\gamma_\nu(\alpha_s) = \alpha_s\,\hat\gamma_{\nu 0} + \alpha_s^2\,\hat\gamma_{\nu 1} + \alpha_s^3\,\gamma_{\nu 2}(\theta_2^{\gamma_\nu})$$

$$H(\alpha_s) = \Big|\hat c_0 + \alpha_s\,\hat c_1 + \alpha_s^2\,c_2(\theta_2^H)\Big|^2$$

$$\tilde S(\alpha_s) = \Big[\hat{\tilde S}_0 + \alpha_s\,\hat{\tilde S}_1 + \alpha_s^2\,\tilde S_2(\theta_2^S)\Big]^2$$

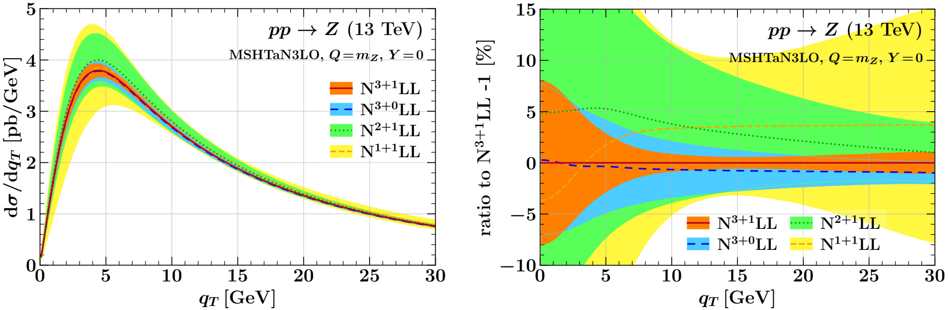
- Up to 5 one-dimensional functional series for beam functions (plus several more for DGLAP splitting functions)

$$\tilde b_i(x, \alpha_s) = \sum_j \int \frac{\mathrm{d}z}{z}\Big[\hat I_{ij,0}(z) + \hat I_{ij,1}(z) + I_{ij,2}(z, \theta_2^{B_{ij}})\Big]\, f_j\Big(\frac{x}{z}\Big)\,,$$

  - Currently use known functional form: $I_{ij,n}(z, \theta_n^{B_{ij}}) = \frac{3}{2}\,\theta_n^{B_{ij}}\,\hat I_{ij,n}(z)$
  - In the future use strategy 2 to parameterize z dependence

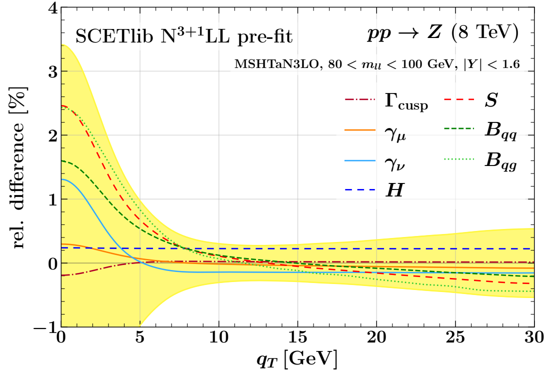
Results for Drell-Yan p_T Spectrum.

Comparing different orders at 95% “theory CL” ($\boldsymbol{\Delta\theta_n = 2}$)



- Uncertainties reduce as we go to higher order (by construction)

Uncertainty Breakdown.

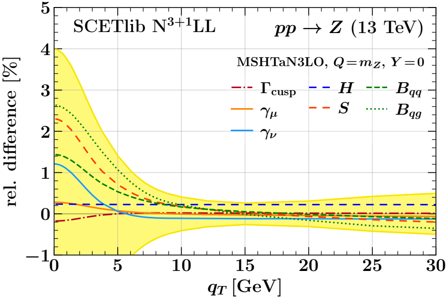


Separately varying each TNP by $\pm\Delta\theta_n = 1$

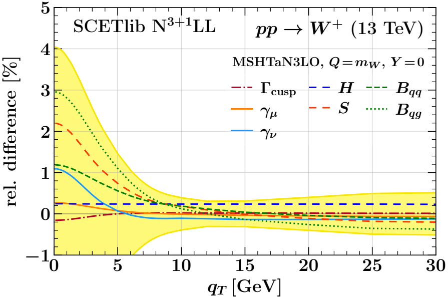
- TNPs provide breakdown into independent uncertainty sources with correct shape
  - Encodes correct point-by-point correlations
  - Importantly, carries over to p_T^ℓ and other decay kinematics

Correlations between W and Z.

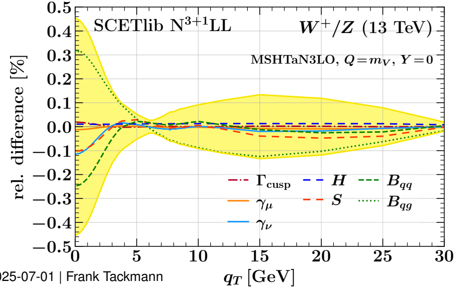
relative impacts for Z



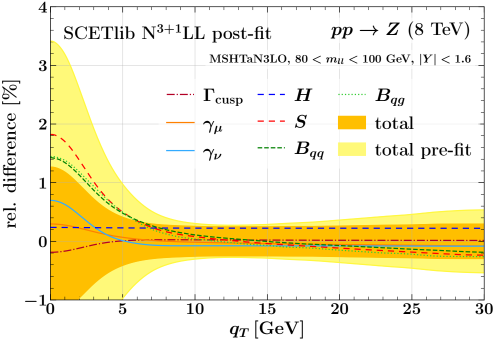
relative impacts for W

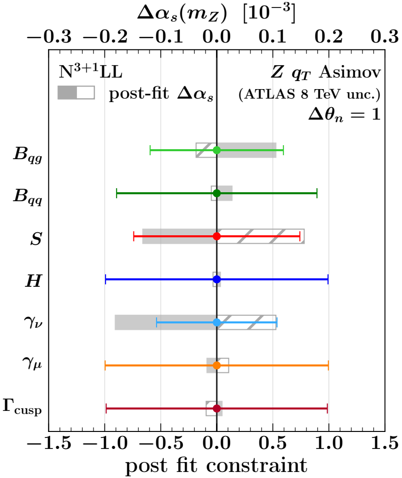


relative impacts on W/Z



- Caveats apply: These are only the (formerly) leading perturbative unc.
- Subleading effects can become important (or even dominant) now
  - Quark mass effects
  - EW corrections
  - Power corrections

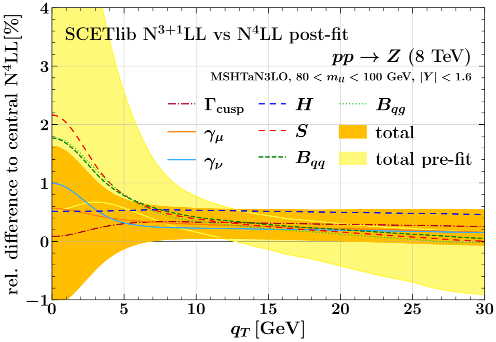


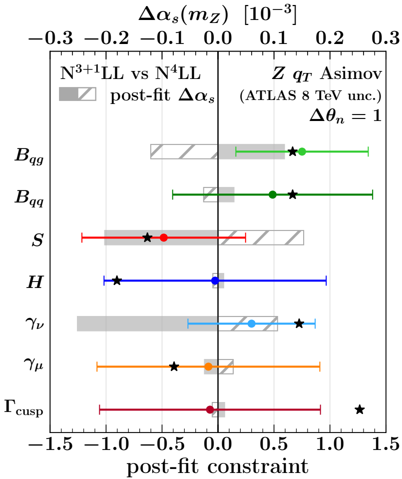


Fit to pseudodata generated from N^{3+1}LL

- Include prior Gaussian theory constraint $\theta_n = 0 \pm 1$
- Data provides nontrivial constraints on TNPs
  - Post-fit prediction has reduced theory uncertainties
  - Induces nontrivial post-fit correlations

Profile against True Higher Order.

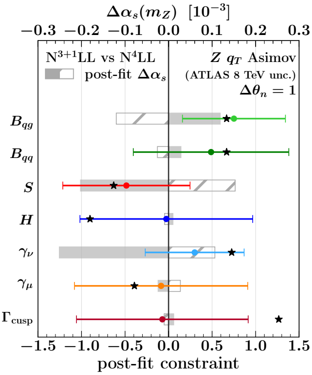




Fit to pseudodata generated from true N^{4}LL

- Simulates fit to real data (which contains all-order result)
  - TNPs are pulled toward their true values
  - Post-fit prediction gets corrected toward true result

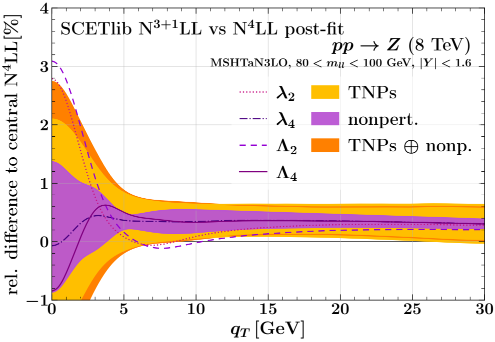
Relaxing the Prior Theory Constraint.

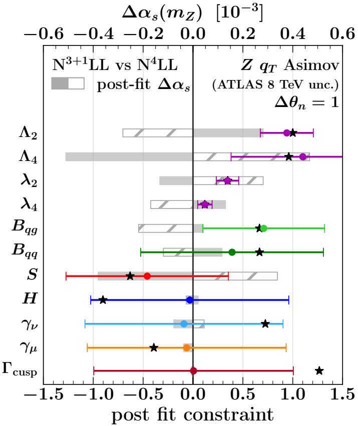


Data is able to sufficiently constrain TNPs by itself

- Reduces dependence on prior theory constraint (and associated potential bias)
  - Post-fit constraints on TNPs become even more consistent with true values
- Uncertainty on final result almost unchanged

Bonus: Including Nonperturbative Effects.





For $1/b_T \sim q_T \gg \Lambda_{\rm QCD}$ nonpert. effects can be systematically expanded in an OPE

$$\tilde{f}_i(x, b_T, \mu, Q) = \tilde{f}_i^{(0)}(x, b_T, \mu, Q)\left\{1 + b_T^2\left[\Lambda_{2,i}(x) + \lambda_2 \ln\frac{b_T Q}{b_0}\right] + \mathcal{O}(\Lambda_{\rm QCD}^4 b_T^4)\right\}$$

- Also include quadratic and quartic OPE coefficients in the fit
  - TNPs still pulled toward their true values but less constrained

Summary.

Interpretation of precision measurements requires *meaningful* theory uncertainties which includes in particular proper theory correlations

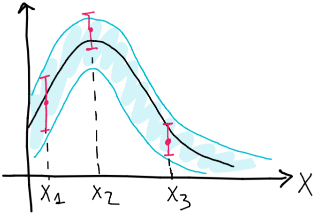
Scale variations become insufficient once theory unc. $\sim$ experimental unc

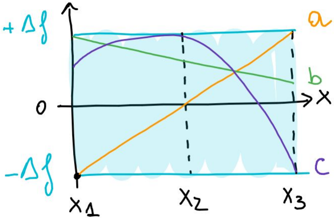
- Neither particularly reliable nor can they do correlations

Theory nuisance parameters

- Provide truly parametric theory uncertainties that
  - ✓ Encode correct correlations
  - ✓ Can be consistently propagated everywhere (fits, MCs, neural networks, ...)
  - ✓ Can be consistently profiled and constrained by data
- Bonus: Can fully benefit from all known partial higher-order information
- First successful applications to resummed Drell-Yan p_T spectrum
  - Implemented in SCETlib (available upon request, hopefully fully public soon)
  - Precision W-mass measurement by CMS [→ see Kenneth's talk]
- First promising applications to PDF fits and fixed-order predictions [MSHT20aN3LO (Gowan et al.) '22; Poncelet, Lim '24]

Additional Slides





	ρ_{12}	ρ_{13}	ρ_{23}
a	0	−1	0
b	+1	+1	+1
c	+1	−1	−1

Repeat fit for each individual scale variation and take envelope of results

- Amounts to trying out various correlation models for the same total uncertainty band
  - None of the trial variations provides a realistic correlation model
  - Individual variations are not meaningful (which is why we take their envelope)
- Best we can do with scale variations
  - Perform as many variations as we can to “fill out” the band, hoping to include at least one that happens to give sufficiently conservative estimate
  - And/or identify conceptually “independent” subsets of variations and add their envelopes

TNPs for Drell-Yan p_T Spectrum: More Details.

Apply strategy 2 with $\varepsilon = q_T^2/Q^2$ ($Q \equiv \sqrt{m_{\ell\ell}}$)

$$\frac{\mathrm{d}\sigma}{\mathrm{d}^4 q} = \frac{\mathrm{d}\sigma^{(0)}}{\mathrm{d}^4 q} \times \left[1 + \mathcal{O}\left(\frac{q_T^2}{Q^2}\right)\right]$$

- Power corrections stay below $\lesssim 5\%$ up to $q_T \lesssim Q/3 \ldots Q/2$
- Leading-power term is subject of q_T factorization and resummation

$$\frac{\mathrm{d}\sigma^{(0)}}{\mathrm{d}^4 q} = \frac{1}{2E_{\mathrm{cm}}^2} L_{VV'}(q^2) \sum_{a,b} H_{VV'\,ab}(q^2, \mu)$$

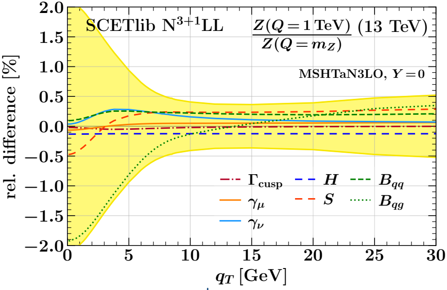
$$\times \int \frac{\mathrm{d}^2 \vec{b}_T}{(2\pi)^2} e^{\mathrm{i}\vec{b}_T \cdot \vec{q}_T}\, \widetilde{B}_a(x_a, b_T, \mu, \nu/Q)\, \widetilde{B}_b(x_b, b_T, \mu, \nu/Q) \tilde{S}(b_T, \mu, \nu)$$

$$x_{a,b} = \frac{Q}{E_{\mathrm{cm}}} e^{\pm Y}$$

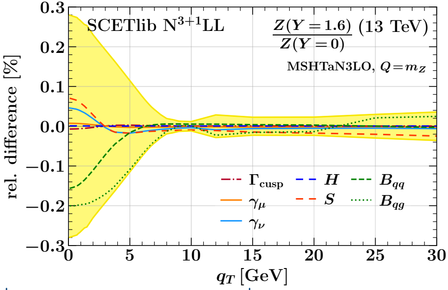
- Here $VV' = \{\gamma\gamma, \gamma Z, Z\gamma, ZZ, W^+W^+, W^-W^-\}$
- Factorization allows us to apply strategy 1 to $q_T(b_T)$, q^2, V
- Also allows us to factorize $x_{a,b}$ dependence and apply strategy 2 to it

Results for Drell-Yan p_T Spectrum: Other Ratios.

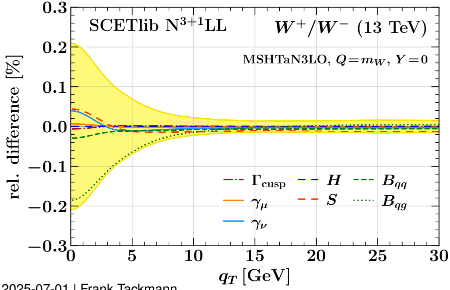
$Z(Q = 1\,\text{TeV})/Z(Q = m_Z)$



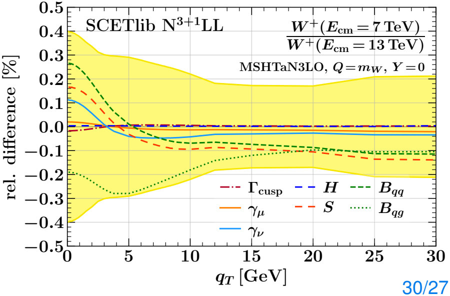
$Z(Y = 1.6)/Z(Y = 0)$



W^+/W^-



$W^+(E_{\text{cm}} = 7\,\text{TeV})/W^+(E_{\text{cm}} = 13\,\text{TeV})$



Asimov Fit Setup.

[For details see Cridge, Marinelli, FT; arXiv:2506.13874]

We perform Asimov fits to (unfluctuated) pseudodata

- Standard method to study expected uncertainties in a controlled setting
  - Unobscured by statistical fluctuations and subleading effects
- Goals: Demonstrate TNPs and estimate expected uncertainties in $\alpha_s(m_Z)$
  - Can consistently drop subleading effects in both pseudodata and theory model (power corrections, quark mass effects, EW corrections)
  - They are needed to fit the real data, but are irrelevant for estimating the dominant uncertainties

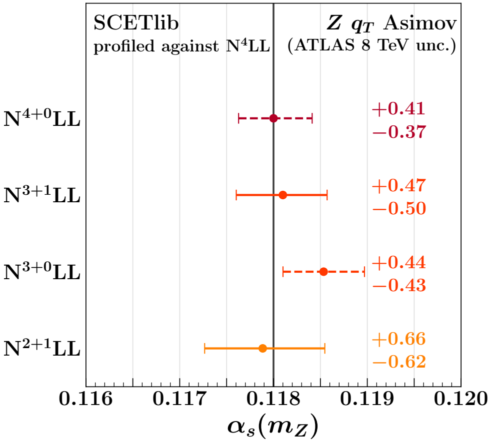
Pseudodata

- Central value given by central SCETlib prediction with $\alpha_s(m_Z) = 0.118$
- Exp. uncertainties and correlations from ATLAS 8 TeV inclusive $Z\ p_T$ measurement [Eur. Phys. J. C 84 (2024) 315 [arXiv: 2309.09318]]
- Same bins and cuts as used by ATLAS $\alpha_s(m_Z)$ determination [arXiv:2309.12986]

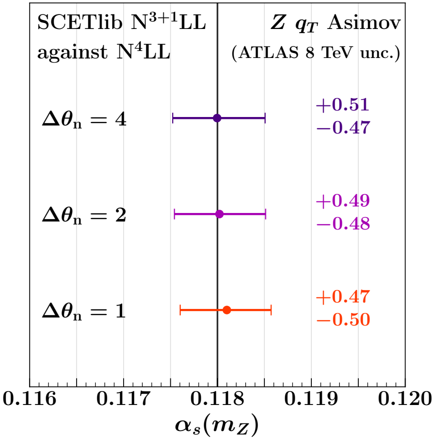
Results for α_s with Profiling TNPs.

[For details see Cridge, Marinelli, FT; arXiv:2506.13874]

different pert. orders



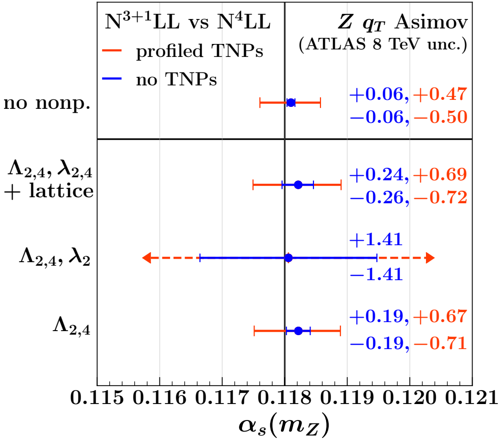
relaxing theory constraint



- Important: Not all sources of theory uncertainties included yet

α_s Results with TNPs and Nonpert. Parameters.

[For details see Cridge, Marinelli, FT; arXiv:2506.13874]



- Important: Not all sources of theory uncertainties included yet

Acknowledgments.

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 101002090 COLORFREE)

European Research Council

Established by the European Commission